

#2

**THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re the Application of : Naoki MATSUOKA, et al.

Filed : Concurrently herewith

For : PACKET SWITCH DEVICE AND....

Serial No. : Concurrently herewith

January 12, 2001

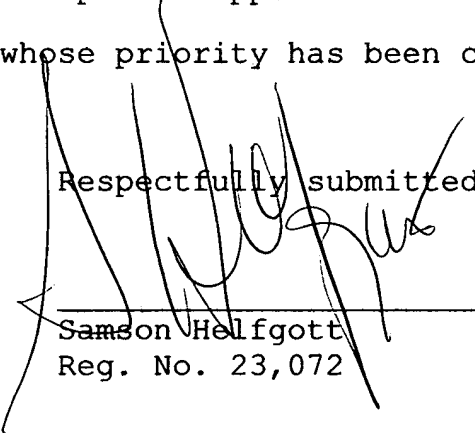
Assistant Commissioner of Patents  
Washington, D.C. 20231

**SUBMISSION OF PRIORITY DOCUMENT**

S I R:

Attached herewith are Japanese patent application No.  
2000-006360 of January 12, 2000 whose priority has been claimed  
in the present application.

Respectfully submitted

  
\_\_\_\_\_  
Samson Helfgott  
Reg. No. 23,072

HELFGOTT & KARAS, P.C.  
60th FLOOR  
EMPIRE STATE BUILDING  
NEW YORK, NY 10118  
DOCKET NO.:FUJY 17.298  
BHU:priority

Filed Via Express Mail

Rec. No.: EL522394436US

On: January 12, 2001

By: Brendy Lynn Belony

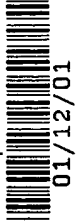
Any fee due as a result of this paper,  
not covered by an enclosed check may be  
charged on Deposit Acct. No. 08-1634.

JC974 U.S. PTO  
09/759181  
01/12/01

日本国特許庁

PATENT OFFICE  
JAPANESE GOVERNMENT

JC974 U.S. PTO  
09/759181



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日

Date of Application:

2000年 1月12日

出願番号

Application Number:

特願2000-006360

出願人

Applicant(s):

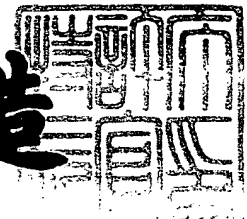
富士通株式会社

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2000年 9月 1日

特許庁長官  
Commissioner,  
Patent Office

及川耕造



出証番号 出証特2000-3070879

【書類名】 特許願

【整理番号】 9951485

【提出日】 平成12年 1月12日

【あて先】 特許庁長官殿

【国際特許分類】 H04L 12/28

【発明の名称】 パケットスイッチ装置及びスケジューリング制御方法

【請求項の数】 12

【発明者】

    【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号富士通株式会社内

    【氏名】 松岡 直樹

【発明者】

    【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号富士通株式会社内

    【氏名】 朝永 博

【発明者】

    【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号富士通株式会社内

    【氏名】 瓦井 健一

【特許出願人】

    【識別番号】 000005223

    【氏名又は名称】 富士通株式会社

【代理人】

    【識別番号】 100089244

    【弁理士】

    【氏名又は名称】 遠山 勉

【選任した代理人】

    【識別番号】 100090516

    【弁理士】

【氏名又は名称】 松倉 秀実

【連絡先】 0 3 - 3 6 6 9 - 6 5 7 1

【手数料の表示】

【予納台帳番号】 012092

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9705606

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 パケットスイッチ装置及びスケジューリング制御方法

【特許請求の範囲】

【請求項 1】 複数の入力回線のいずれかに対応して設けられ、論理的に複数の出力回線対応のキューに分割されたバッファメモリをそれぞれ有し、対応の前記入力回線を通して入力されたパケットを固定長パケットの形態で前記バッファメモリに一時的に蓄積する複数の入力バッファ部と；

前記複数の入力バッファ部から送出された前記固定長パケットを前記複数の出力回線のいずれかに送出するためのスイッチング動作を行うパケットスイッチ部と；

前記入力バッファ部からの前記固定長パケットの送出スケジューリング処理を前記出力回線の数に対応する複数の単位時間をかけ、かつ前記送出スケジューリング処理を前記入力回線の数に対応する複数並列で実施するパイプライン・スケジューリング処理手段と、一つのフレームを構成する前記固定長パケットの送出状態を前記入力回線毎に管理する送出状態管理手段とを有し、前記複数の出力回線のいずれかに対応して設けられる複数のスケジューラ部と；

前記複数のスケジューラ部のそれぞれにおける前記送出スケジューリング処理の結果情報を対応する前記入力バッファ部に通知する少なくとも一つの結果通知部とを備え；

前記複数並列の送出スケジューリング処理において同一フレーム対応の前記固定長パケットを送出中の前記入力回線を選択せず、選択確定後は同一フレーム対応の前記固定長パケットの送出完了まで同一入力回線を選択を維持するパケットスイッチ装置。

【請求項 2】 前記スケジューラ部のそれぞれは、前記入力バッファ部に蓄積されている前記固定長パケットの送出要求情報を前記入力回線毎に管理する要求管理手段を更に有し、

前記パイプライン・スケジューリング処理手段は、前記要求管理手段からの前記送出要求情報と、前記スケジューラ部間を環状に接続する伝送媒体を通して入力される選択未確定情報と、前記スケジューラ部を相互に接続する伝送媒体を通

して前記送出状態管理手段に入力される送出状態情報とに基づいて、前記固定長パケットの送出を行う前記入力回線対応の前記入力バッファ部を決定する請求項 1 記載のパケットスイッチ装置。

【請求項 3】 前記結果通知部は、前記入力バッファ部が一つのフレームを構成する前記固定長パケットを送出中であるか否かを前記送出スケジューリング処理の結果情報に基づいて管理する管理手段と、

前記複数のスケジューラ部のそれぞれにおける前記結果情報を対応する前記入力バッファ部に通知する場合、一つのフレームを構成する前記固定長パケットの送出中に異なる前記スケジューラ部から入力された前記結果情報を廃棄する廃棄処理手段と、

を有する請求項 1 記載のパケットスイッチ装置。

【請求項 4】 前記スケジューラ部及び前記結果通知部を前記複数の出力回線のいずれかに対応させて設け、かつ前記複数の出力回線の増加に応じて拡張可能に縦続接続する手段を備える請求項 1 記載のパケットスイッチ装置。

【請求項 5】 前記固定長パケットは、一つのフレームを構成し複数に分割された第 1 のパケットと、それぞれが一つのフレームを構成する第 2 のパケットとを含み、

前記第 2 のパケットのそれぞれにフレームエンド識別子を付加する手段を前記入力バッファ部の前段に備える請求項 1 記載のパケットスイッチ装置。

【請求項 6】 前記パイプライン・スケジューリング処理手段は、前記入力バッファ部から送出された送出要求情報を QoS クラス毎に管理し、前記出力回線における QoS 帯域制御及び QoS 優先制御結果に基づいて選択される QoS クラスに対して、前記送出スケジューリング処理を行う請求項 1 記載のパケットスイッチ装置。

【請求項 7】 前記スケジューラ部のそれぞれは、所定時間内に前記入力バッファ部から入力された送出要求情報の数を計数する負荷観測手段と、

この負荷観測の結果を前記パイプライン・スケジューリング処理手段に通知する契機を前記入力バッファ部毎に一単位時間ずつシフトさせて指示する観測結果反映手段とを更に有し、

観測した前記送出要求情報の数の割合に応じて前記入力回線対応の前記入力バッファ部を選択する請求項 1 記載の packetswitch 装置。

【請求項 8】 可変長データのフレームを複数の固定長 packetswitch 単位に変換して前記固定長 packetswitch 単位で、入力回線から出力回線にスイッチングする packetswitch 装置において；

複数の入力回線に対してそれぞれ設けられる複数の入力バッファ部と；

同一の前記フレームから生成された前記固定長 packetswitch を同一の出力回線に送出完了するまで連続して読み出しできるようにスケジューリングするスケジューラとを設け；

前記スケジューラの指示に従って各入力バッファ部から前記固定長 packetswitch を読み出してスイッチングする packetswitch 装置。

【請求項 9】 複数の入力回線のいずれかに対応して設けられ、論理的に複数の出力回線対応のキューに分割されたバッファメモリをそれぞれ有し、対応の前記入力回線を通して入力された packetswitch を固定長 packetswitch の形態で前記バッファメモリに一時的に蓄積する複数の入力バッファ部からの、前記固定長 packetswitch の送出スケジューリング処理を前記出力回線の数に対応する複数の単位時間をかけ、かつ前記送出スケジューリング処理を前記入力回線の数に対応する複数並列のパイプライン処理で実施する手順と；

一つのフレームを構成する前記固定長 packetswitch の送出状態を前記入力回線毎に管理する手順と；

前記送出スケジューリング処理の結果情報を対応する前記入力バッファ部に通知する手順と；

前記複数並列の送出スケジューリング処理において同一フレーム対応の前記固定長 packetswitch を送出中の前記入力回線を選択せず、選択確定後は同一フレーム対応の前記固定長 packetswitch の送出完了まで同一入力回線を選択を維持する手順と；  
を備えるスケジューリング制御方法。

【請求項 10】 前記入力バッファ部から送出された送出要求情報を QoS クラス毎に管理し、前記出力回線における QoS 帯域制御及び QoS 優先制御結果に基づいて選択される QoS クラスに対して、前記送出スケジューリング処理を

行う手順を更に備える請求項 9 記載のスケジューリング制御方法。

【請求項 1 1】 所定時間内に前記入力バッファ部から入力された送出要求情報の数を計数する手順と；

この負荷観測の結果を前記送出スケジューリング処理に反映させる契機を前記入力バッファ部毎に一単位時間ずつシフトさせて指示する手順と；

観測した前記送出要求情報の数の割合に応じて前記入力バッファ部を選択する手順とを更に備える請求項 9 記載のスケジューリング制御方法。

【請求項 1 2】 可変長データのフレームを複数の固定長パケットに変換して前記固定長パケット単位で、入力回線から出力回線にスイッチングするパケットスイッチ装置において；

複数の入力回線から入力する前記固定長パケットを各入力回線ごとに設けられる入力バッファに蓄積する手順と；

同一の前記フレームから生成された前記固定長パケットを同一の出力回線に送出完了するまで連続して読み出しできるようにスケジューリングする手順と；

前記スケジューリングで指示された前記入力バッファから前記固定長パケットを読み出す手順と；

を備えるスケジューリング制御方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は広帯域交換機、クロスコネクトスイッチ装置、及びルータ装置などに適用され、固定長パケットを伝送（特に、限定しないときは、交換、伝達及び転送を含む）するパケットスイッチ装置に関し、特に大規模パケットスイッチ装置の一構成法である入力バッファ型パケットスイッチ装置に関する。

【0002】

本発明は、更に詳細には、セルと称される固定長パケットのスイッチングを行うパケットスイッチ装置（ATMスイッチ装置）において、IP（Internet Protocol）パケットのような可変長パケットが入力されたときでも効率的に伝送することを可能にするパケットスイッチ装置に関する。



【 0 0 0 3 】

【従来の技術】

近年のインターネット網の普及に伴い、I P 通信ネットワークシステムにおける I P データ（音声情報を含む）のトラヒックが急激に増加しており、数年先には、I P トラヒックが通信トラヒックの大多数を占めると考えられている。

【 0 0 0 4 】

この背景の一要因としては、ローカルエリアネットワーク（L A N）の広帯域化等をはじめとするネットワーク技術の進歩と、パーソナルコンピュータ（P C）の多機能化及び P C に適用する C P U の高速化等をはじめとする P C 技術の進歩とにともない、複数の L A N 上の P C 間でデータだけではなく音声情報をも高速に通信することが実用的に可能になってきたことがある。

【 0 0 0 5 】

これにより、従来の電話による音声の通信を専用線、L A N 及びワイドエリアネットワーク（W A N）などから構成されるインターネット網上の P C 間で実行するアプリケーションソフトウェアと、このソフトウェアを組み込んだハードウェアシステムとが急速に市場に投入されている。このシステムは、電話網とインターネット網などの I P パケット交換網とを統合することにより、設備及び運用などのコストを削減できる複合交換網システム、つまり I P 通信ネットワークシステムである。

【 0 0 0 6 】

上述したような通信ネットワークインフラの多様化・拡大状況の中、企業は、増え続ける通信トラヒックに対処するために、大容量でかつ効率的に I P パケットを伝送できる大規模パケットスイッチ装置を実現することを推進している。

【 0 0 0 7 】

I P パケットは可変長のパケット形態で伝送されるが、厳密な可変長データのスイッチングはバイト単位に処理が必要なため、一般的には高速処理が必要であり、高速スイッチングが難しい。

【 0 0 0 8 】

そこで、可変長パケットを高速にスイッチングする手法として、可変長パケッ

トをセルと同様に 5 3 バイト長の固定長パケットに分割し、パケットスイッチ装置内では固定長パケットのスイッチングを行っている。しかし、パケットスイッチ装置からデータを送出する際に、分割された固定長パケットを分割前の元の可変長パケットに戻してパケットを伝送しなければならない。

## 【 0 0 0 9 】

ここで、従来のパケットスイッチ装置において、可変長パケットを固定長パケットに分割して転送する 2 つの手法について、図 1 を参照して説明する。

(1) スケジューラ部 (図示省略) は、入力バッファ部を有する入力インタフェース部 (入力 INF) において、フレーム (複数の固定長パケットで構成される可変長データ) を意識せず、分割された固定長パケット毎にスケジューリングを行い、共通スイッチ部としてのパケットスイッチ (SW) に入力する。また、パケットスイッチ (SW) の後段の出力インタフェース部 (出力 INF) に設けられたフレーム組み立てのための出力バッファ部において、待ち合わせを行ってフレームの組み立てを行う (図 1 (A) 参照)。

## 【 0 0 1 0 】

(2) スケジューラ部 (図示省略) は、入力バッファ部を有する入力インタフェース部 (入力 INF) において、フレームを意識し、同一フレームを構成するパケットを連続してスケジューリングし、パケットスイッチ (SW) に入力する。また、パケットスイッチ (SW) でフレーム単位にスイッチングを行った後、出力インタフェース部 (出力 INF) を通して出力回線 (出力方路) に送出する (図 1 (B) 参照)。

## 【 0 0 1 1 】

## 【発明が解決しようとする課題】

上記 (1) の手法では、スケジューリングの際にフレームを意識しないため、固定長パケット毎にスケジューリングを実施する。そのため、スケジューリング処理を並列化させ、先行して次の周期のスケジューリングを行うパイプライン処理の適用が可能であり、処理の高速化を図れるという利点を持つ。その反面、出力インタフェース部において、フレーム単位にパケット送出を行うために、各入力回線 (入力方路) からの固定長パケットを一旦蓄積し、フレームの組立処理が

必要である。

【 0 0 1 2 】

この手法を採用したパケットスイッチ装置においては、出力インタフェース部毎に入力回線数対応のフレーム組み立てのためのバッファメモリが必要となる。従って、スイッチ規模の大きな装置では、多数の出力バッファメモリが必要である。

【 0 0 1 3 】

一方、上記（２）の手法では、フレーム単位のスイッチングにより、出力インタフェース部のフレーム組み立てのための出力バッファ部が不要になる利点を持つ。しかし、フレーム毎のスケジューリングを実現するために、フレーム送出中の入力回線に対して、他の出力回線がスケジューリングを行ってはならず、また一旦確定した入力回線はフレーム送出完了まで継続してスケジューリングする必要がある。

【 0 0 1 4 】

従って、スケジューリングの際に、現在どの入力回線がどの出力回線に対してフレーム送出中であるかという最新の情報を参照する必要があり、上述したパイプライン処理のようにスケジューリング処理を並列化させ、先行して次の周期のスケジューリングを行うことができない。この結果、一単位時間内に全回線のスケジューリングを実施する高速なスケジューラ部が必要となり、大規模パケットスイッチ装置での実現が困難である。

【 0 0 1 5 】

本発明の課題は、固定長パケットのスイッチングを行うパケットスイッチ装置において、ＩＰパケットのような可変長パケットが入力されたときでも効率的に伝送することを可能にするパケットスイッチ装置及びスケジューリング制御方法を提供することにある。

【 0 0 1 6 】

本発明の他の課題は、パイプライン処理により必要な処理速度の低減を図りつつ、フレーム単位のスケジューリング処理を可能にするパケットスイッチ装置及びスケジューリング制御方法を提供することにある。

## 【 0 0 1 7 】

本発明の別の課題は、フレーム組立のための出力バッファ部を必要としないパケットスイッチ装置及びスケジューリング制御方法を提供することにある。

## 【 0 0 1 8 】

## 【課題を解決するための手段】

上記課題を解決するために、本発明の第1のパケットスイッチ装置は、複数の入力回線のいずれかに対応して設けられ、論理的に複数の出力回線対応のキューに分割されたバッファメモリをそれぞれ有し、対応の前記入力回線を通して入力されたパケットを固定長パケットの形態で前記バッファメモリに一時的に蓄積する複数の入力バッファ部と；

前記複数の入力バッファ部から送出された前記固定長パケットを前記複数の出力回線のいずれかに送出するためのスイッチング動作を行うパケットスイッチ部と；

前記入力バッファ部からの前記固定長パケットの送出スケジューリング処理を前記出力回線の数に対応する複数の単位時間をかけ、かつ前記送出スケジューリング処理を前記入力回線の数に対応する複数並列で実施するパイプライン・スケジューリング処理手段と、一つのフレームを構成する前記固定長パケットの送出状態を前記入力回線毎に管理する送出状態管理手段とを有し、前記複数の出力回線のいずれかに対応して設けられる複数のスケジューラ部と；

前記複数のスケジューラ部のそれぞれにおける前記送出スケジューリング処理の結果情報を対応する前記入力バッファ部に通知する少なくとも一つの結果通知部とを備え；

前記複数並列の送出スケジューリング処理において同一フレーム対応の前記固定長パケットを送出中の前記入力回線を選択せず、選択確定後は同一フレーム対応の前記固定長パケットの送出完了まで同一入力回線を選択を維持する。

## 【 0 0 1 9 】

本発明の第2のパケットスイッチ装置は、上記第1のパケットスイッチ装置において、前記スケジューラ部のそれぞれは、前記入力バッファ部に蓄積されている前記固定長パケットの送出要求情報を前記入力回線毎に管理する要求管理手段

を更に有し、

前記パイプライン・スケジューリング処理手段は、前記要求管理手段からの前記送出要求情報と、前記スケジューラ部間を環状に接続する伝送媒体を通して入力される選択未確定情報と、前記スケジューラ部を相互に接続する伝送媒体を通して前記送出状態管理手段に入力される送出状態情報とに基づいて、前記固定長パケットの送出を行う前記入力回線対応の前記入力バッファ部を決定する。

【 0 0 2 0 】

本発明の第 1 及び第 2 のパケットスイッチ装置においては、スケジューリング処理を入力回線数に対応して複数 (N) 並列で実行し、かつ全スケジューラ部共通のフレーム送出状態が設定された送出状態管理手段を参照して、自己スケジューラ部がフレーム送出中の場合は、継続して同一入力回線を選択し、フレーム送出中でない場合には、フレーム送出中の入力回線を選択しないようにしている。その結果、一つのフレームを構成する固定長パケットを連続して読み出すことが可能となる。また、並列処理により処理速度の低減が可能となる。

【 0 0 2 1 】

本発明の第 3 のパケットスイッチ装置は、上記第 1 のパケットスイッチ装置において、前記結果通知部は、前記入力バッファ部が一つのフレームを構成する前記固定長パケットを送出中であるか否かを前記送出スケジューリング処理の結果情報に基づいて管理する管理手段と、

前記複数のスケジューラ部のそれぞれにおける前記結果情報を対応する前記入力バッファ部に通知する場合、一つのフレームを構成する前記固定長パケットの送出中に異なる前記スケジューラ部から入力された前記結果情報を廃棄する廃棄処理手段とを有する。

【 0 0 2 2 】

本発明の第 3 のパケットスイッチ装置においては、N 個の並列処理において、複数のスケジューラ部が同時に同一入力回線に対してスケジューリングを行った場合に、結果通知部 (結果切り替え部) において、入力バッファ部に通知されるスケジューリング結果が 1 つだけになるように、後から通知されるスケジューリング結果を廃棄するようにしている。従って、各並列処理において競合が発生し

た場合にも、フレーム単位の読み出しを維持することができる。

【 0 0 2 3 】

本発明の第 4 のパケットスイッチ装置は、上記第 1 のパケットスイッチ装置において、前記スケジューラ部及び前記結果通知部を前記複数の出力回線のいずれかに対応させて設け、かつ前記複数の出力回線の増加に応じて拡張可能に縦続接続する手段を備える。

【 0 0 2 4 】

本発明の第 4 のパケットスイッチ装置においては、スケジューラ部と結果通知部とを出力回線毎に分散的に配備するようにしている。よって、スイッチ規模の拡張時に、新たなスケジューラ部と結果通知部とを追加することによって、オンラインでのスケジューラの拡張が可能になる。

【 0 0 2 5 】

本発明の第 5 のパケットスイッチ装置は、上記第 1 のパケットスイッチ装置において、前記固定長パケットは、一つのフレームを構成し複数の分割された第 1 のパケットと、それぞれが一つのフレームを構成する第 2 のパケットとを含み、

前記第 2 のパケットのそれぞれにフレームエンド識別子を付加する手段を前記入力バッファ部の前段に備える。

【 0 0 2 6 】

本発明の第 5 のパケットスイッチ装置においては、A T M セル等のようにフレーム化を必要としない固定長パケットに対して、フレームの最終パケットを示す情報を付与するようにしている。従って、スケジューラ部では、単位フレーム長のパケットとして扱われるため、固定長パケットと可変長パケットを意識することなく、一元的に処理することが可能となる。

【 0 0 2 7 】

本発明の第 6 のパケットスイッチ装置は、上記第 1 のパケットスイッチ装置において、前記パイプライン・スケジューリング処理手段は、前記入力バッファ部から送出された送出要求情報を Q o S クラス毎に管理し、前記出力回線における Q o S 帯域制御及び Q o S 優先制御結果に基づいて選択される Q o S クラスに対して、前記送出スケジューリング処理を行う。

【 0 0 2 8 】

本発明の第 6 のパケットスイッチ装置においては、入力バッファ部からの送出要求情報を入力回線毎かつ Q o S クラス毎に管理するようにしている。そして、帯域、優先制御等の Q o S 制御により選択された Q o S クラスについてスケジューリングを行うようにしている。従って、入力バッファ部において、パケットスイッチの出力回線における Q o S 制御を実現することができ、品質劣化ポイントを入力バッファ部のみに限定することができ、呼受付制御等の制御が容易に実現できる。

【 0 0 2 9 】

本発明の第 7 のパケットスイッチ装置は、上記第 1 のパケットスイッチ装置において、前記スケジューラ部のそれぞれは、所定時間内に前記入力バッファ部から入力された送出要求情報の数を計数する負荷観測手段と、

この負荷観測の結果を前記パイプライン・スケジューリング処理手段に通知する契機を前記入力バッファ部毎に一単位時間ずつシフトさせて指示する観測結果反映手段とを更に有し、

観測した前記送出要求情報の数の割合に応じて前記入力回線対応の前記入力バッファ部を選択する。

【 0 0 3 0 】

本発明の第 7 のパケットスイッチ装置においては、入力バッファ部から送出された送出要求情報を入力回線毎に計数する負荷観測手段を備え、計数した要求数を一定周期毎にパイプライン・スケジューリング処理手段に反映する。この反映する契機を入力回線毎に一単位時間ずつ遅らせて反映させるようにしている。これにより、反映契機が重ならないため、入力回線数の多い場合でも容易に処理することが可能となる。

【 0 0 3 1 】

本発明の第 8 のパケットスイッチ装置は、可変長データのフレームを複数の固定長パケットに変換して前記固定長パケット単位で、入力回線から出力回線にスイッチングするパケットスイッチ装置において；

複数の入力回線に対してそれぞれ設けられる複数の入力バッファ部と；

同一の前記フレームから生成された前記固定長パケットを同一の出力回線に送出完了するまで連続して読み出しできるようにスケジューリングするスケジューラとを設け；

前記スケジューラの指示に従って各入力バッファ部から前記固定長パケットを読み出してスイッチングする。

#### 【 0 0 3 2 】

本発明の第 8 のパケットスイッチ装置においては、一つのフレームを構成する固定長パケットを連続して読み出すことが可能となる。

本発明の第 1 のスケジューリング制御方法は、複数の入力回線のいずれかに対応して設けられ、論理的に複数の出力回線対応のキューに分割されたバッファメモリをそれぞれ有し、対応の前記入力回線を通して入力されたパケットを固定長パケットの形態で前記バッファメモリに一時的に蓄積する複数の入力バッファ部からの、前記固定長パケットの送出スケジューリング処理を前記出力回線の数に対応する複数の単位時間をかけ、かつ前記送出スケジューリング処理を前記入力回線の数に対応する複数並列のパイプライン処理で実施する手順と；

一つのフレームを構成する前記固定長パケットの送出状態を前記入力回線毎に管理する手順と；

前記送出スケジューリング処理の結果情報を対応する前記入力バッファ部に通知する手順と；

前記複数並列の送出スケジューリング処理において同一フレーム対応の前記固定長パケットを送出中の前記入力回線を選択せず、選択確定後は同一フレーム対応の前記固定長パケットの送出完了まで同一入力回線を選択を維持する手順とを備える。

#### 【 0 0 3 3 】

本発明の第 1 のスケジューリング制御方法においては、上記第 1 のパケットスイッチ装置と同様の効果を期待できる。

本発明の第 2 のスケジューリング制御方法は、上記第 1 のスケジューリング制御方法において、前記入力バッファ部から送出された送出要求情報を Q o S クラス毎に管理し、前記出力回線における Q o S 帯域制御及び Q o S 優先制御結果に基



づいて選択されるQoSクラスに対して、前記送出スケジューリング処理を行う手順を更に備える。

【0034】

本発明の第2のスケジューリング制御方法においては、上記第6の PACKET スイッチ装置と同様の効果を期待できる。

本発明の第3のスケジューリング制御方法は、上記第1のスケジューリング制御方法において、所定時間内に前記入力バッファ部から入力された送出要求情報の数を計数する手順と；

この負荷観測の結果を前記送出スケジューリング処理に反映させる契機を前記入力バッファ部毎に一単位時間ずつシフトさせて指示する手順と；

観測した前記送出要求情報の数の割合に応じて前記入力バッファ部を選択する手順とを更に備える。

【0035】

本発明の第3のスケジューリング制御方法においては、上記第7の PACKET スイッチ装置と同様の効果を期待できる。

本発明の第4のスケジューリング制御方法は、可変長データのフレームを複数の固定長 PACKET に変換して前記固定長 PACKET 単位で、入力回線から出力回線にスイッチングする PACKET スイッチ装置において；

複数の入力回線から入力する前記固定長 PACKET を各入力回線ごとに設けられる入力バッファに蓄積する手順と；

同一の前記フレームから生成された前記固定長 PACKET を同一の出力回線に送出完了するまで連続して読み出しできるようにスケジューリングする手順と；

前記スケジューリングで指示された前記入力バッファから前記固定長 PACKET を読み出す手順とを備える。

【0036】

本発明の第4のスケジューリング制御方法においては、上記第8の PACKET スイッチ装置と同様の効果を期待できる。

【0037】

【発明の実施の形態】

次に、本発明の実施の形態について図面を参照して説明する。

〔入力バッファ型パケットスイッチ装置〕

（全体構成）

本発明の一実施の形態におけるパケットスイッチ装置の全体構成を示す図 2 を参照すると、入力バッファ型パケットスイッチ装置は、複数（N 個）の入力バッファ部 1，入力 N×出力 M・マトリックスのパケットスイッチ（共通スイッチ）2，複数（M 個）のスケジューラ部 3，及び結果切り替え部 4 を備える。

【0038】

ここで、N 個の入力バッファ部 1 は複数（N 個）の入力回線（入力方路）対応の N 個の入力インタフェース部（図 5 中の入力 I N F）にそれぞれ設けられる。M 個のスケジューラ部 3 及び結果切り替え部 4 は、スケジューラ（スケジューリング制御装置）5 を構成する。N×M・マトリックスのパケットスイッチ 2 からの複数（M 個）の出力回線（出力方路）は、出力インタフェース部（図 5 中の出力 I N F）に收容される。

【0039】

入力バッファ部 1 のそれぞれは、対応の入力回線からの固定長パケット（以下、特に限定する必要がないときは、単に「パケット」と記載する）を M 個の出力回線に対応した論理キュー（F I F O のバッファメモリ）毎にバッファリングする。コアスイッチとしてのパケットスイッチ 2 は入力バッファ部 1 から読み出したパケットをパケットヘッダのルーティングビットに従ってスイッチングする。

【0040】

スケジューラ部 3 のそれぞれは入力バッファ部 1 からの送出要求（送出要求情報）を受けて、パケットスイッチ 2 の出力回線において競合が発生しないように読み出しを制御する。結果切り替え部 4 は出力回線毎のスケジューラ部 3 から受信したスケジューリング結果（スケジューリング結果情報）を対応する入力バッファ部 1 へ通知する。出力インタフェース部は低速回線への論理多重分離（DMUX）機能を有する。

【0041】

上述した入力バッファ型パケットスイッチ装置においては、各入力バッファ部

1 に出力回線毎の論理キューを設けることにより、入力バッファ型パケットスイッチ装置構成の基本的特徴である低いメモリアクセス速度を実現しつつ、HOL (Head of Line) ブロッキングを回避できる。また、メモリアクセス速度が低いので、汎用のRAMを入力バッファ部1に適用可能であり、大容量のバッファメモリが実現できる。

#### 【0042】

##### (スケジューラ部の基本構成)

上記スケジューラ部3のそれぞれは、図3に示すような基本構成を採る。M個の出力回線に対応してそれぞれ設けられるM個のスケジューラ部3は同一構成である。各スケジューラ部3において、要求管理部31は入力バッファ部1からの送出要求情報を各入力回線(#0…#N-1)毎に管理する。N個のスケジューリング処理部(#0…#N-1)32は自己対応の出力回線のスケジューリング処理をN並列で処理する。フレーム送出状態管理部33はフレーム送出中の入力回線番号とその出力先(現在、フレーム送出中の出力回線番号)とを管理する。

#### 【0043】

スケジューリング結果通知部34はスケジューリング処理部32において確定されたスケジューリング結果情報を各シーケンス(パイプライン処理)毎に保持し、所定の時刻に保持していたスケジューリング結果情報(入力回線番号、スケジューラ番号を含む)を結果切り替え部4に送出する。なお、図3中、DISは分配部、SELはセレクタである。

#### 【0044】

スケジューリング処理部32は、スケジューラ部間を環状に接続するリンクを通して入力される未確定情報と、スケジューラ部3を相互に接続するバスを通してフレーム送出状態管理部33に入力されるフレーム送出状態情報と、入力バッファ部1からの送出要求情報とに基づいて、パケットの送出を行う入力回線を決定する。

#### 【0045】

##### (結果切り替え部の基本構成)

上記結果切り替え部 4 は図 4 に示すような基本構成を採る。結果切り替え部 4 において、切り替えスイッチ部 4 1 はスケジューラ部 3 から入力された入力回線番号及びスケジューラ番号を含むスケジューリング結果情報を、結果を通知すべき入力バッファ部 1 の方路へスイッチングする。フレーム送出管理部 4 2 は入力バッファ部 1 の数に対応して N 個設けられる。各フレーム送出管理部 4 2 は、対応の入力バッファ部 1 がフレーム送出中であることを管理する。

## 【 0 0 4 6 】

廃棄処理部 4 3 は入力バッファ部 1 の数に対応して N 個設けられる。各廃棄処理部 4 3 はフレーム送出管理部 4 2 の指示に従いスケジューリング結果情報を廃棄する。結果切り替え部 4 は、スケジューラ部 3 からのスケジューリング結果情報を各入力バッファ部 1 へ通知すると同時に、フレーム送出中に異なるスケジューラ部 3 から要求が入力された場合、その結果を廃棄して競合制御を行う。

## 【 0 0 4 7 】

## 〔スケジューリング処理〕

図 2 及び図 5 を併せ参照して、一つのフレームを構成する複数の固定長パケットを連続して読み出す制御について説明する。

## 【 0 0 4 8 】

入力バッファ型パケットスイッチ装置の入力インタフェース部（入力 I N F）に設けられる入力バッファ部 1 のそれぞれは、論理的に出力回線（出力方路）毎に分かれた複数の論理キュー（F I F O のバッファメモリ）を有する。スケジューラ部 3 がフレームを意識せずにスケジューリングを行った場合、図 5（A）に示すように、パケットスイッチ（S W）2 の一つの出力回線上では、他の入力回線から到着したパケットによって、フレーム構成が崩れてしまう。

## 【 0 0 4 9 】

フレーム構成を崩さないためには、図 5（B）に示すように、一つの入力回線からのパケットをある一つの出力回線に対して送出し始めた場合、それ以降必ず同一入力回線からのパケットをその出力回線に対して送出できるように、スケジューリングを行う必要がある。

## 【 0 0 5 0 】

このため、スケジューラ部 3 は、一度確定した入力回線に対して、その入力回線からの一フレーム分のパケットを送出完了するまで継続してスケジューリングを行うとともに、未確定の状態のときには、既にフレームを構成するパケットを送出中（以下、フレーム送中と記載する）の入力回線を選択しないようにする。

#### 【 0 0 5 1 】

次に、具体的なスケジューリング処理手法について説明する。ここでは、説明の簡略化のため、4 入力・4 出力構成の入力バッファ型パケットスイッチ装置 1 に基づいて説明する。

#### 【 0 0 5 2 】

図 2 及び図 3 を併せ参照すると、スケジューラ部 3 の要求管理部 3 1 は、入力バッファ部 1 から入力された（到着した）要求情報中の出力回線番号が自己スケジューラ番号と一致した場合に、入力回線番号（要求情報の入力ポート番号から判断可能）に対応したカウンタをインクリメント（+フレーム長）し、またスケジューリング処理により、確定した入力回線に対応するカウンタをデクリメント（-1）することで、要求数を管理している。

#### 【 0 0 5 3 】

そして、要求管理部 3 1 は、この要求数が 1 以上の入力回線をスケジューリング要求ありと判定する。要求情報は出力回線番号、QoS クラス、及びフレーム長からなり、入力バッファ部 1 にフレームの最後の固定長パケットが到着した時に通知される。

#### 【 0 0 5 4 】

この要求管理部 3 1 の動作フローを示す図 6 を参照すると、要求管理部 3 1 は、入力バッファ部 1 からの要求情報の到着イベントか、スケジューラ部 3 のスケジューリング終了イベントであるかを判定する（処理手順 S 6 1）。次に、入力バッファ部 1 からの要求情報到着の場合、到着情報が自己スケジューラに対するものか否かを、要求情報の出力回線番号 Out-No. と予め設定される自己スケジューラ番号 SCH-No. との比較により判定する（S 6 2）。

#### 【 0 0 5 5 】

自己スケジューラ宛で無い場合は処理を終了し、自己スケジューラ宛の場合は、要求情報の入力回線番号に対応する要求管理カウンタ  $Req. cnt$  をインクリメント（+フレーム長）する（S63）。次に、要求数が1以上を示す入力回線毎のフラグをセットする、つまり要求数有無表示レジスタ  $Req. reg$  に“1”を設定する（S64）。

#### 【0056】

一方、スケジューラ部3のスケジューリング終了イベント発生時には（S61）、確定した入力回線に対応する要求管理カウンタ  $Req. reg$  をデクリメント（-1）する（S65）。要求管理カウンタ  $Req. reg$  の内容の“0”判定を行い、カウンタ値が“1”以上であれば処理を終了し、“0”の場合には、要求数の有無を示すフラグをクリアする、つまり要求数有無表示レジスタ  $Req. reg$  に“0”を設定する（S66, S67）。

#### 【0057】

次に、スケジューラ部3のスケジューリング処理部32は、図7中の上部に示すタイムスロット図のように、1回のスケジューリング処理を4単位時間かけて、この処理を4パラレル処理で実現する。スケジューリング対象となるスケジューラ番号及びパイプラインシーケンスは、全スケジューラ共通の処理シーケンス（図7上部のタイムスロット図）に従って与えられる。この例では、時刻  $T = a$  において、出力回線#2のパイプラインシーケンス#0、出力回線#1のパイプラインシーケンス#1、出力回線#4のパイプラインシーケンス#2、及び出力回線#3のパイプラインシーケンス#3の処理が並列処理で実行される。

#### 【0058】

スケジューリング指示を受けたスケジューラ部3は、未確定情報、要求情報、及びフレーム送出状態情報を参照してスケジューリングを行う。これらの情報はすべて各入力回線に対応したビットマップレジスタに格納される。

#### 【0059】

未確定情報は、図7中、「A」と表記されたレジスタに格納され、“0”，“1”によってその入力回線が未確定であるか既に確定済みであることを表す。また、要求情報は、図7中、「B」と表記されたレジスタに格納され、要求管理部3

1の要求数が1以上の入力回線に対して”1”が設定され、要求数の有無を表す。フレーム送出状態情報は、図7中、「C」と表記されたレジスタに格納され、フレーム送出状態管理部33において設定され、各入力回線がフレーム送出中（”1”）であるかを示す。

#### 【0060】

次に、スケジューラ部3のフレーム送出状態管理部33は、各入力回線に対応した上記フレーム送出状態を示すビットマップのレジスタ（初期値0）を有しており、スケジューリングにより確定した入力回線に対するこのレジスタのビットが”0”のとき（”0”はフレームの先頭パケットに対するスケジューリングを意味する）、レジスタに”1”を設定する。そして、フレームの最終パケットに関するスケジューリングが行われたときに、このビットをクリアする。このレジスタにより、”1”の設定されている入力回線がフレーム送出中であることを知ることができる。

#### 【0061】

フレームの最終パケットに関するスケジューリングであるか否かは、図8に示す処理により知ることができる。この処理はフレーム送出状態管理部33において実施され、各入力回線に対応したスタートポイントSP及びエンドポイントEPと、フレーム長が設定されたリンクテーブルとを有する。

#### 【0062】

このリンクテーブルはフレーム長カウンタの役割を果たし、スタートポイントSP及びエンドポイントEPは各入力回線のリンクテーブルの先頭アドレス及び最終アドレスをそれぞれ示す。入力回線毎にフレーム長をリンクテーブルで管理し、スケジューリング確定時にこのフレーム長の値を減算してゆき、フレーム長=0になったとき、フレームの最終パケットに対するスケジューリングと見なす。

#### 【0063】

一層具体的に動作例を説明すると、入力回線#0には、フレーム長「2→5→3」の要求情報が、かつ入力回線#N-1には、フレーム長「3→5」の要求情報が既に到着しており、これらの情報がリンクテーブルに設定されているとする

## 【 0 0 6 4 】

ここでは、スケジューリング処理の結果、入力回線 # 0 が確定したとする。この場合、対応する入力回線のスタートポインタ S P を参照し、スタートポインタ S P が示す値 " a " をリンクテーブルのアドレスとして、フレーム長情報を参照して更新 ( - 1 減算 ) を行う ( 図 8 ( A ) 参照 ) 。

## 【 0 0 6 5 】

次のスケジューリング周期で再び入力回線 # 0 が選択され、上記処理と同様にリンクテーブルの減算を行うとフレーム長が " 0 " になる。このフレーム長 " 0 " をもって最終パケットのスケジューリングであるか否かを判定し、入力回線 # 0 のフレーム送出状態レジスタをクリアする。そして、入力回線 # 0 のスタートポインタ S P を N e x t のフィールドに設定された値に更新する ( 図 8 ( B ) 参照 ) 。

## 【 0 0 6 6 】

また、新たに要求情報が到着した場合には、エンドポインタ E P が示すアドレスにフレーム長と次のアドレスとを書き込み、エンドポインタ E P を更新する ( 図 8 ( C ) 参照 ) 。上記処理を繰り返すことにより、フレームの最終パケットに対するスケジューリングであるか否かを判定する ( 図 8 ( D ) 参照 ) 。

## 【 0 0 6 7 】

なお、上述した動作例では、図 8 に示す処理をスケジューラ部 3 のフレーム送出状態管理部 3 3 に行わせているが、入力バッファ部 1 で実施させてもよい。その場合、入力バッファ部 1 からパケットが送出されたときに上記処理を行い、フレーム最終パケットと判定されたとき、その旨を要求情報と一緒にスケジューラ部 3 のフレーム送出状態管理部 3 3 に通知する。

## 【 0 0 6 8 】

次に、図 7 を参照してフレーム単位のスケジューリング動作を説明する。時刻  $T = a$  において、全ての入力回線が未確定であり、スケジューラ部 # 1 及びスケジューラ部 # 2 は全ての入力回線から送出要求を受けており、全ての入力回線はフレーム未送出状態であるとする。



## 【 0 0 6 9 】

スケジューリングアルゴリズムとしては、S R R (Sequential Round Robin) 方式を採用することができる。この S R R 方式は各入力バッファ部間に優先度を与え、優先度の高い入力バッファ部においてどの出力回線に対応する論理キューから読み出すかを、入力負荷に応じたラウンドロビン制御により決定していく手法である。入力バッファ部の優先度はタイムスロット毎に決まった順序でシーケンシャルに与える。なお、このスケジューリングアルゴリズムの詳細については、特願平 1 0 - 3 5 5 8 8 8 号 (H 1 0 . 1 2 . 1 5 ) や、特願平 1 0 - 2 4 5 3 3 1 号 (H 1 0 . 8 . 3 1 ) など参照できる。

## 【 0 0 7 0 】

このスケジューリングアルゴリズムに基づいて、要求情報があった未確定の入力回線で、かつフレーム未送出の入力回線の中からひとつが選択される。つまり、ポインタが示す入力回線からサーチして上記条件を満足する入力回線で最初に検出されたものを選択する。

## 【 0 0 7 1 】

図 7 に示す例では、スケジューラ部 # 2 が入力回線 # 3 を選択したとする。その結果、入力回線 # 3 対応のフレーム送出状態情報のレジスタ C 及び未確定情報のレジスタ A にそれぞれ " 1 " が設定され、スケジューリングが確定した入力回線番号を保持しておく。この保持レジスタを以降スケジューリングレジスタと称する。また、この処理と並行してスケジューラ部 # 1 では、同様の手順で入力回線 # 2 を確定し、フレーム送出状態情報のレジスタ C、未確定情報のレジスタ A 及びスケジューリングレジスタをそれぞれ更新する。

## 【 0 0 7 2 】

次のスケジューリング処理の際、スケジューリングレジスタが設定されている場合は、設定されている入力回線を選択する。また、設定されていない場合は、フレーム送出状態情報レジスタ C が " 0 " のフレーム未送出の入力回線であり、未確定、かつ要求がある入力回線の中からスケジューリングを行う。

## 【 0 0 7 3 】

この例では、前スケジューリング処理において、スケジューリングレジスタが

設定されているため、次のスケジューリング時は、スケジューリングレジスタに設定された入力回線、つまりスケジューラ部 # 2 は入力回線 # 3、スケジューラ部 # 1 は入力回線 # 2 を選択する。なお、フレーム送出状態管理部 3 3 より、最終フレームである旨の通知を受けたとき、このスケジューリングレジスタはクリアされる。

## 【 0 0 7 4 】

当然のことながら、スケジューラ部 # 4 及びスケジューラ部 # 3 のスケジューリング処理では、フレーム送出状態情報が設定されている入力回線 # 2 及び入力回線 # 3 を選択することはできない。

## 【 0 0 7 5 】

このように、上記処理により、一つの入力回線について一フレーム連続してスケジューリングを行い、またフレーム送出中の入力回線を選択しないようにしているため、フレーム単位での読み出しを実現するとともに、高いスループット性能を得ることができる。

## 【 0 0 7 6 】

図 9 はスケジューリング処理フローを示す。スケジューリング処理においては、まずフレーム送出（送出中）フラグがセットされているか否かを判定する（処理手順 S 9 1）。初期のフラグはリセット状態であるため、S 9 2 において、スケジューリングモードの判定処理を行う。

## 【 0 0 7 7 】

負荷観測モード（負荷読み出しモード）でない場合は、未確定、要求あり、かつフレーム未送出状態の入力回線の中から一つの入力回線を選択する（S 9 4）。負荷観測モードの場合は、未確定、要求あり、フレーム未送出状態、かつ累積カウンタ（図 1 9 中の L D. r e g に対応）の内容が 1 以上の状態の入力回線の中から一つの入力回線を選択する（S 9 5）。

## 【 0 0 7 8 】

次に、S 9 6 において、選択した確定入力回線をフレーム送出中の入力回線として保持し、S 9 7 において、フレーム送出フラグをセットする。続いて、S 9 8 において、フレームエンドか否かの判定を行う。フレームエンドの場合は、フ

レーム送出フラグをクリアして処理を終了する（S 9 9）。

【 0 0 7 9 】

一方、フレームエンドで無い場合は、次のスケジューリング処理において、再び S 9 1 の判定を行う。ここでは、既にフレーム送出フラグがセットされているため、S 9 3 において、S 9 6 で保持しておいた入力回線を選択する。続いて、S 9 7 でフレーム送出フラグをセットする処理の後、S 9 8 でフレームエンド判定を行い、以降、フレームエンドになるまで上記処理を繰り返す。

【 0 0 8 0 】

なお、上記処理によってスケジューリングされた結果は、各パイプラインシーケンス毎に保持され、各パイプラインシーケンス終了時（図 7 中、太線が記されたスロット）、結果切り替え部 4 に通知され、確定された入力バッファ部に対して読み出し指示の結果情報が通知される。

【 0 0 8 1 】

上述したスケジューリング処理では、同一時刻に確定したスケジューリング結果が異なっていたが、同じ入力回線に対してスケジューリングを行ってしまった場合の処理について、図 1 0 を参照して説明する。

【 0 0 8 2 】

通常は、未確定情報を参照しているため、異なる出力回線対応のスケジューラ部 3 が同一入力回線を選択することは無い。しかし、上述したように未確定の状態においては、各スケジューラ部に対して全ての入力回線が未確定として通知されるため、スケジューリング処理によっては、同一入力回線を選択することが起こり得る。

【 0 0 8 3 】

図 1 0 に示す例では、時刻  $T = b$  において、スケジューラ部 # 2 及びスケジューラ部 # 1 が同一入力回線 #  $N - 1$  に対してスケジューリングを行っている。このスケジューリング結果は、上述したようにパイプライン毎に保持され、パイプライン処理の最後に結果切り替え部 4 に通知される。従って、パイプラインシーケンス # 0 の結果は時刻  $T = d$  で、かつパイプラインシーケンス # 1 の結果は時刻  $T = e$  で通知されることになる。すなわち、スケジューラ部 # 2 のスケジュー

リング結果がスケジューラ部 # 1 の結果より先に結果切り替え部 4 に通知される。

#### 【 0 0 8 4 】

結果切り替え部 4 では、各スケジューラ部 3 のフレーム送出状態管理部 3 3 （図 3 参照）と同様のフレーム送出管理部 4 2 （図 4 参照）を備えており、各入力バッファ部 1 がフレーム送出中であるか否かを以下の制御により管理している。

#### 【 0 0 8 5 】

つまり、結果切り替え部 4 では、新たにスケジューラ部 3 からスケジューリング結果情報が到着したとき、フレーム送出管理用レジスタに " 1 " をセットし、フレームの最終パケットに対する結果が到着したときにクリアする。結果切り替え部 4 は、" 1 " が設定されている間をフレーム送出中とみなす。結果切り替え部 4 では、上記動作例の様にフレーム長管理制御は行わなくてもよい。その場合、スケジューラ部 3 において判定した結果（フレームの最終パケットに関するスケジューリングであるか否か）をスケジューリング結果と一緒に結果切り替え部 4 に通知し、結果切り替え部 4 ではその情報を基に上記フレーム送出状態の管理を行う。

#### 【 0 0 8 6 】

結果切り替え部 4 では、時刻  $T = d$  において、スケジューラ部 # 2 の結果が到着したことにより、入力回線 #  $N - 1$  のフレーム送出管理用レジスタはフレーム送出中フラグ ( " 1 " ) がセットされると同時に、スケジューラ番号 ( # 2 ) が保持される。そして、時刻  $T = e$  において、入力回線 #  $N - 1$  に対するスケジューリング結果がスケジューラ部 # 1 から到着したとき、フレーム送出状態情報とスケジューラ番号とのチェックを行う。フレーム送出管理部 #  $N - 1$  はフレーム送出中 ( " 1 " ) で、かつ保持したスケジューラ番号と到着したスケジューラ番号とが異なる場合に、廃棄処理部 4 3 （図 4 参照）に対して到着結果（スケジューリング結果）情報を廃棄するよう指示を行う。

#### 【 0 0 8 7 】

この例では、保持されたスケジューラ番号が # 2 で、かつ到着したスケジューラ番号が # 1 であるため、このスケジューラ部 # 1 から来たスケジューリング結

果情報は廃棄されることになる。廃棄を行った場合には、結果切り替え部 4 はスケジューラ部 3 に対して廃棄された旨を通知する。このとき、スケジューラ部 3 ではスケジューリングが完了したときに行う処理（要求管理数のデクリメント等）を実行しないようにする。

#### 【 0 0 8 8 】

このように、同一時刻に同一入力回線に対してスケジューリングが行われた場合にも、最初に入力バッファ部 1 に通知されたスケジューリング結果以外は入力バッファ部 1 に通知されないため、フレーム毎の読み出しが可能になる。

#### 〔スケジューラの拡張構成〕

次に、スケジューラ 5 の拡張構成法について、図 1 1 を参照して説明する。スケジューラ部 3 はパケットスイッチの出力回線毎に設けることが可能である。このとき、入力バッファ部 1 とスケジューラ部 3 間の情報（要求通知及びスケジューリング結果通知）が膨大になることを回避するため、各入力バッファ部 1 からの要求情報（Req # 0 … Req # N - 1）をひとつのスケジューラ部 3 で受信し、その情報を他のスケジューラ部 3 へ通知する構成を採る。また、スケジューリング結果情報（Ack # 0 … Ack # N - 1）も各スケジューラ部 3 から個々に通知するのではなく、結果切り替え部 4 を経由して一括して通知するようにする。

#### 【 0 0 8 9 】

出力回線対応の各スケジューラ部 3 はスケジューリング要求情報の入力ハイウェイ（HW）と拡張出力ハイウェイ（HW）とを備える。そして、受信したスケジューリング要求情報をそのまま拡張出力ハイウェイへ出力することで、各スケジューラ部 3 へ要求情報を通知する。さらに、スケジューリング結果を入力バッファ部 1 へ通知する結果切り替え部 4 を出力回線毎に対応させて設け、1 入力ハイウェイと N 出力ハイウェイ及び N 拡張入力ハイウェイとからなる 1 x N 構成の結果切り替え部 4 とする。

#### 【 0 0 9 0 】

拡張時はスケジューラ部 3 と結果切り替え部 4 とのセット（同図中、点線で囲んだ部分）を追加することによってスケジューラ（スケジューリング制御装置）

5の拡張が可能となる。さらに、運用中においても現在運用されている部分を止めることなく随時拡張が可能であり、インサービスでの増設が可能である。

#### 【 0 0 9 1 】

上記拡張例では、スケジューラ部3に拡張出力ハイウェイを設けているが、図12に示すように、各スケジューラ部3に同一の要求情報を分岐するバス構成を採ってもよい。

#### 【 0 0 9 2 】

〔可変長パケット及び固定長パケットの混在収容〕

次に、可変長パケット及び固定長パケットを入力バッファ型パケットスイッチ装置に混在収容する構成例について、図13を参照して説明する。例えば、ATM（非同期転送モード：Asynchronous Transfer Mode）セルなどの固定長パケットはパケットスイッチ装置から出力する際にフレームを構成して出力させる必要はない。このようなフレーム化が不必要なサービスの固定長パケットと、IPパケットの様にフレーム化が必要なサービスの可変長パケットとを混在収容する場合について説明する。

#### 【 0 0 9 3 】

インタフェース（入力インタフェース部）毎に収容サービスが異なる場合は、ATMセルを収容するフレーム化の必要の無いインタフェースでは、全てのパケットに対してフレームの最終パケットを意味する識別子を付与するパケットタイプ付与部6を入力インタフェース部などに設ける。この識別子は、パケットのヘッダ部分の空き領域を使用するか、スケジューラ部に通知するための別経路を設けることにより容易に実現される。その結果、入力バッファ部1あるいはスケジューラ5では、固定長パケットをフレーム長=1の可変長パケットと見なして処理することができ、サービスを意識してスケジューリング処理を変える必要はない。

#### 【 0 0 9 4 】

可変長パケットはパケットタイプ付与部6の前段で固定長に分割され、各分割パケットには装置内ヘッダが付与されるとともに、一つのフレームを構成する最終パケットだけにはパケットタイプ付与部6によってフレームエンド識別子が付

与される。よって、同一スイッチ内に様々なサービスのデータパケットを混在収容することができる。

#### 【 0 0 9 5 】

なお、上述した例では、インタフェース種別毎に全てフレームエンド識別子を付与するか否かを決定するようにしているが、同一インタフェースに複数のサービス形態のパケットが収容される場合には、そのサービス種別を識別し、そのサービス毎に全てのパケットのエンド識別子を付けるか否かを決定してもよい。ここで、入力バッファ型パケットスイッチ装置の他の各構成要素は、図 2 に示す構成または上述した各構成例を採ることができる。

#### 【 0 0 9 6 】

〔 Q o S 制御機能を有するスケジューラ部 〕

次に、Q o S 制御構成法について、図 1 4，図 1 5 及び図 1 6 を参照して説明する。

#### 【 0 0 9 7 】

一般的には、図 1 4 ( A ) に示すように、入力バッファ型パケットスイッチ装置の出力回線において Q o S が保証されるように Q o S 制御を実施する。Q o S 制御は、Q o S クラス毎に予め決められた帯域に出力データをシェーピング（帯域制限）することで、Best Effort サービスから帯域、遅延を重視する Guaranteed サービスクラスの品質を保証したり、遅延を重視する Q o S クラスを優先的に送出する優先制御等によって、各 Q o S クラスの品質を保証することを目的としている。

#### 【 0 0 9 8 】

入力バッファ型パケットスイッチ装置では、基本的には出力側にバッファ部を有していないため、入力バッファ部からパケットを読み出す際に、予め出力回線の Q o S 制御を行った上で、スケジューリング処理を行うことにより、スイッチ装置の出力回線における Q o S 保証を実現する。

#### 【 0 0 9 9 】

ここでは、スケジューラ（スケジューリング制御装置）において Q o S 制御を行う場合について説明する。図 1 5 を参照すると、スケジューラのスケジューラ

部 3 は、要求管理部 3 1 において、入力バッファ部 1 から受信した要求情報を入力回線毎かつ Q o S クラス毎に管理する。Q o S 制御部 3 5 は、上述したように予め設定された帯域を満足しており、かつ優先制御の優先順序に従って送出 Q o S クラスを決定する。

【 0 1 0 0 】

スケジューリング処理部 3 2 では、指定された Q o S クラスの要求情報をもとにスケジューリング処理を実施する。例えば、出力回線 # a では、Q o S クラス # 0 が最優先で、次に Q o S クラス # 1 が高い優先クラスとする。Q o S クラス # 0 が最優先であるため、Q o S クラス # 0 に要求があれば、スケジューリング対象 Q o S クラスを Q o S # 0 と決定する。Q o S クラス # 0 に要求が無ければ Q o S # 1 をスケジューリング対象 Q o S クラスとする。

【 0 1 0 1 】

帯域制御を行う場合は、設定された帯域（＝ 1 / 帯域の間隔）で読み出し指示を与える回路（例えば、リーキバケット回路）を用いて、各 Q o S クラスが送出間隔を満たしているかをチェックし、送出間隔を満たしている Q o S クラスの中で、優先度の高い Q o S クラスをスケジューリング対象 Q o S クラスとする。

【 0 1 0 2 】

このように、入力バッファ部 1 からのパケット読み出し時に、出力回線における Q o S 制御結果を反映させてスケジューリングを行うことにより、出力回線側にバッファ部が無くても出力回線における Q o S 保証を実現することができ、品質劣化点を入力バッファ部 1 のみにすることができる。これは、呼受付制御等の処理が入力バッファ部 1 の一箇所だけで済むため、呼受付処理などの制御が容易になるメリットがある。

【 0 1 0 3 】

また、図 1 4 （ B ）に示すように、パケットスイッチ装置の出力側、つまり出力インタフェース部（出力 I N F ）に低速回線への多重分離部（ D M U X ）がある場合も上記と同様の制御で実現できる。

【 0 1 0 4 】

この場合、スケジューラ部 3 の要求管理部 3 1 において、 D M U X の出力ポー



ト ( $\#0 \cdots \#j-1$ ) 毎、かつ入力回線 ( $\#0 \cdots \#N-1$ ) 毎及び QoS クラス ( $QoS \#0 \cdots \#q-1$ ) 毎に要求情報を管理する。要求管理部 31 は出力ポートの数に対応して  $j$  個設けられる。そして、スケジューリングを行う前に、まずスケジューリングを行う DMUX ポートを決定する。例えば、 $\text{mod. } 4$  の DMUX であれば、4 単位時間を 1 周期として、ポート 1 → ポート 2 → ポート 3 → ポート 4 → ポート 1 → ポート 2 …… という順序で DMUX ポートが選択される。

## 【0105】

スケジューリング処理部 32 は、指定されたポート番号 ( $\#0 \cdots \#j-1$ ) で、かつ QoS 制御部 35 によって指示された QoS クラス ( $QoS \#0 \cdots \#q-1$ ) の要求情報を用いてスケジューリングを行うことで、低速回線収容時にも QoS 保証を実現することができる。

## 【0106】

図 16 に QoS 制御フローを示す。処理手順 S1602, S1602, S1603 は、QoS クラス毎のリーキバケット処理を示している。S1601 において、各 QoS クラスのリーキバケットカウンタ (LBC) に加算値  $a$  (ここでは、 $+1$ ) を加え、S1602, S1603 により QoS クラスを変えて QoS クラス分の処理を行う。

## 【0107】

S1604 では、LBC の値が設定された閾値 ( $TH=0$ ) 以上であるかを判定する。この閾値以上の状態は、一定のレートで読み出し可能であることを示し、S1605 において読み出し可能 QoS としてフラグを "1" にセットし、LBC をデクリメント ( $-b$ ) する。この LBC の減算値  $b$  は回線レートをその QoS のレートで除算して求められる。

## 【0108】

LBC が閾値未満の場合は、S1610 においてフラグを "0" にセットする。そして、S1607 においてフラグ判定を行い、フラグが "1" に設定されていれば、S1611 においてスケジューリング処理を行い、そうでなければ S1608 において QoS クラス ( $q$ ) を変えて、再び S1604 から処理を繰り返す。

## 【 0 1 0 9 】

なお、S 1 6 0 6 及び S 1 6 0 9 は、いずれの Q o S クラスも閾値未満の場合に処理を終了させるための判定ステップであり、全 Q o S クラス分 (Q) の回数の処理を行ったのち、処理の終了に遷移する。

## 【 0 1 1 0 】

一方、S 1 6 0 7 におけるフラグ判定により、フラグが " 1 " と判定された場合には、S 1 6 1 1 においてこのフラグのセットされた Q o S クラスについてスケジューリング処理を行う。そして、S 1 6 1 2 において上記スケジューリング処理によって確定されたか否かを判定する。なお、確定時には処理を終了し、未確定時には、再び S 1 6 1 3 を経由して S 1 6 0 4 から処理を繰り返す。S 1 6 1 2 において  $j > Q$  の判定を行っているが、これはどの Q o S クラスも未確定状態の場合に処理を終了するためのステップである。

## 【 0 1 1 1 】

〔負荷観測機能を有するスケジューラ部〕

次に、負荷観測機能を有するスケジューラ部について、図 1 7, 図 1 8, 図 1 9 を参照して説明する。複数の入力回線からある出力回線に対して、不均等な負荷でパケットが到着した場合に、各入力回線を均等に選択すると、高負荷入力回線のスループットが低下することが知られている。このスループット低下を回避するために負荷に応じて確定すべき入力回線を決定する。

## 【 0 1 1 2 】

負荷観測機能を有するスケジューラ部の構成例を示す図 1 7 を参照すると、スケジューラ部 3 の負荷観測部 3 6 は、所定時間内に入力バッファ部から到着した要求情報を入力回線毎、かつ Q o S クラス毎にカウントする。そして、負荷観測部 3 6 は所定時間経過後、観測結果を負荷読出制御部 3 7 に反映する。負荷読出制御部 3 7 は、一定周期毎に累積カウンタ（後に詳述）に観測結果を加算し、スケジューリング時に - 1 減算するカウンタ（後に詳述）を入力回線毎、かつ Q o S クラス毎に有している。

## 【 0 1 1 3 】

従って、上記反映契機が全入力回線共通の場合には、一斉に累積カウンタの加

算処理を行うため、処理速度の関係から累積カウンタを独立（並列）に用意する必要がある。

【0 1 1 4】

観測結果反映指示部 3 8 は、各入力回線毎に一単位時間ずつ反映契機をずらして負荷観測部 3 6 に指示を与える。この制御により、各入力回線の反映契機が重なることがなく、累積カウンタをハードウェア規模の少ないメモリで構成することが可能になる。

【0 1 1 5】

続いて、動作例を示す図 1 8 を参照すると、負荷観測カウンタは、要求情報を受信したときインクリメントを行い、スケジューリングが確定したときに累積カウンタが” 0 ” のときに限りデクリメント（- 1）するカウンタである。累積カウンタは、一定周期毎に観測結果を加算し、スケジューリング確定時にデクリメントされるカウンタであり、所定時間内の要求数（負荷）を管理する。

【0 1 1 6】

上記反映契機は、結果反映フレームとして表記されており、反映フレームのハイレベル” H ” の位置で、負荷観測カウンタ及び累積カウンタを加算し、累積カウンタの更新、負荷観測カウンタのリセット、及びレジスタへの反映を行う。このレジスタは、各入力回線に対応したビットマップレジスタであり、累積カウンタの値が” 1 ” 以上のとき、” 1 ” がセットされる。

【0 1 1 7】

すなわち、所定時間内に到着した要求数（厳密には、前観測周期で残った要求数も含む）分スケジューリングされたとき、レジスタが” 0 ” になる。従って、スケジューリングの際に、このレジスタに” 1 ” が設定されている入力回線の中から選択するようにすることで、所定時間内に到着した要求数以上の回数でスケジューリングを行わないように制御することができる。このレジスタは図 9 における負荷 R E G（= 1）に対応する。つまり、負荷（所定時間内に到着した要求数）に応じてスケジューリング回数を制御することができる。

【0 1 1 8】

図 1 8 では、観測周期内に符号 b 1 で示すパケットが 1 個到着しているため、

次の観測周期では前周期に到着した全ての要求情報分スケジューリングされるまで、入力回線 # b のパケットは 1 個以上はスケジューリングされない。2 個目のパケット b 2 がスケジューリングされているのは、ひとつ前の時刻で前周期に到着した全ての要求数がスケジューリングされたからである。上記負荷観測カウンタ、累積カウンタ、及びレジスタは負荷観測部 3 6 に設けられる。

#### 【 0 1 1 9 】

図 1 9 に負荷観測制御フローを示す。処理手順 S 1 9 0 1 … S 1 9 0 5 では、所定時間が経過したか否かを判定する。S 1 9 0 1 において、パケット時間毎に入力回線毎の負荷観測反映フレームカウンタ L T をインクリメントする。S 1 9 0 2 において負荷観測カウンタ L T が設定された負荷観測周期の値 T 1 と等しいか否かを判定する。なお、本処理は、一連の処理を入力回線数 (N) 分繰り返して処理を終了する。

#### 【 0 1 2 0 】

続いて、負荷観測制御 S 1 9 0 6 … S 1 9 1 6 に関する処理を説明する。S 1 9 0 7 において、入力バッファ部 1 からの要求情報到着イベントであるか、スケジューラ部 3 によるスケジューリング確定イベントであるかを判定する。要求情報が到着した場合は、S 1 9 0 8 において負荷観測カウンタ o b s . c n t に到着した要求情報のフレーム数情報を加算することにより、所定時間内に到着した要求数を計測する。

#### 【 0 1 2 1 】

この負荷観測カウンタ o b s . c n t の値は、S 1 9 0 6 において所定時間経過したと判定されたとき、S 1 9 1 4 において、累積カウンタ L D . c n t に加算され、負荷観測カウンタ o b s . c n t をクリアする。そして、S 1 9 1 5 において累積カウンタ L D . c n t が " 0 " でなければ、S 1 9 1 6 において累積カウンタ L D . c n t が " 1 " 以上を示すレジスタ L D . r e g をセットする。なお、このレジスタ L D . r e g はスケジューリング処理の負荷観測モードにおいて使用される。

#### 【 0 1 2 2 】

一方、S 1 9 0 6 においてスケジューラ部 3 のスケジューリング確定イベント

と判定された場合には、まず、S 1 9 0 9において累積カウンタLD. cntが” 0 ” であるかを判定する。累積カウンタLD. cntが” 1 ” 以上の場合は、S 1 9 1 0において累積カウンタLD. cntをデクリメント（- 1）し、累積カウンタLD. cntが” 0 ” の場合は、S 1 9 1 3において負荷観測カウンタobs. cntをデクリメントする。

#### 【 0 1 2 3 】

なお、累積カウンタLD. cntをデクリメントした場合には、S 1 9 1 1において累積カウンタLD. cntの” 0 ” 判定を行い、カウンタ値が” 0 ” になったとき、S 1 9 1 2において累積カウンタLD. cntが” 1 ” 以上を示すレジスタLD. regをクリアする。

#### 【 0 1 2 4 】

#### 【発明の効果】

以上説明したように、本発明によれば、フレーム単位のスケジューリングを並列処理によって実施することにより、出力回線側のフレーム組み立てのためのバッファ部を削除できるだけでなく、処理速度の低減効果を期待できる。この結果、大容量パケットスイッチ装置において、IPパケットなどの可変長パケットの効率的な収容が実現できる。

#### 【図面の簡単な説明】

【図 1】 従来のパケットスイッチ装置の構成を示すブロック図。

【図 2】 本発明の一実施の形態のパケットスイッチ装置の構成を示すブロック図。

【図 3】 図 2 におけるスケジューラ部の基本構成を示すブロック図。

【図 4】 図 2 における結果切り替え部の基本構成を示すブロック図。

【図 5】 フレーム単位のスケジューリングを説明するための図。

【図 6】 要求管理部の処理手順を説明するためのフローチャート。

【図 7】 パイプラインスケジューリング処理を説明するための図。

【図 8】 スケジューラ部におけるフレーム長の管理手法を説明するための図。

【図 9】 スケジューリング処理手順を説明するためのフローチャート。

- 【図 1 0】 競合時の廃棄制御を説明するための図。
- 【図 1 1】 スケジューラの拡張構成の一例を示すブロック図。
- 【図 1 2】 スケジューラの拡張構成の他の例を示すブロック図。
- 【図 1 3】 可変長パケット及び固定長パケットの混在収容を説明するための図。
- 【図 1 4】 Q o S 制御を説明するための図。
- 【図 1 5】 Q o S 制御機能を有するスケジューラ部の構成を示すブロック図。
- 【図 1 6】 Q o S 制御の処理手順を説明するためのフローチャート。
- 【図 1 7】 負荷観測機能を有するスケジューラ部の構成を示すブロック図。
- 【図 1 8】 負荷観測部の動作を説明するための図。
- 【図 1 9】 負荷観測制御の処理手順を説明するためのフローチャート。

【符号の説明】

- 1 入力バッファ部
- 2 パケットスイッチ
- 3 スケジューラ部
- 4 結果切り替え部
- 5 スケジューラ
- 6 パケットタイプ付与部
- 3 1 要求管理部
- 3 2 スケジューリング処理部
- 3 3 フレーム送出状態管理部
- 3 4 スケジューリング結果通知部
- 3 5 Q o S 制御部
- 3 6 負荷観測部
- 3 7 負荷読出制御部
- 3 8 観測結果反映指示部
- 4 1 切り替えスイッチ部
- 4 2 フレーム送出管理部

特 2 0 0 0 - 0 0 6 3 6 0

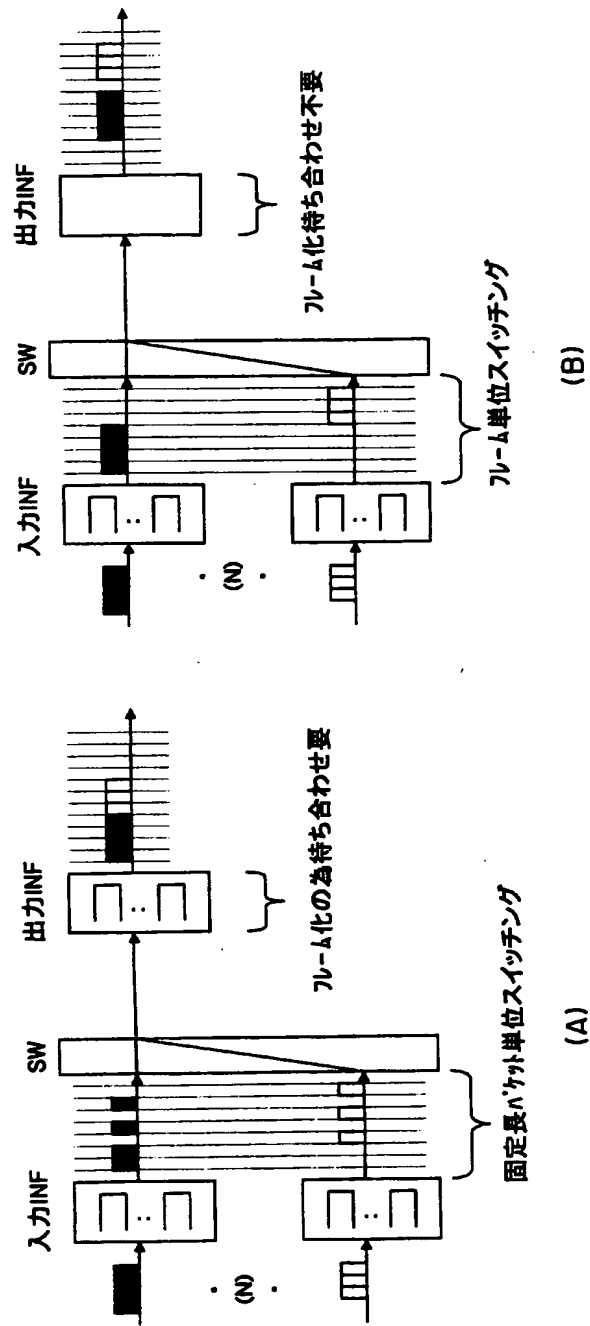
4 3 廃棄処理部

【書類名】

図面

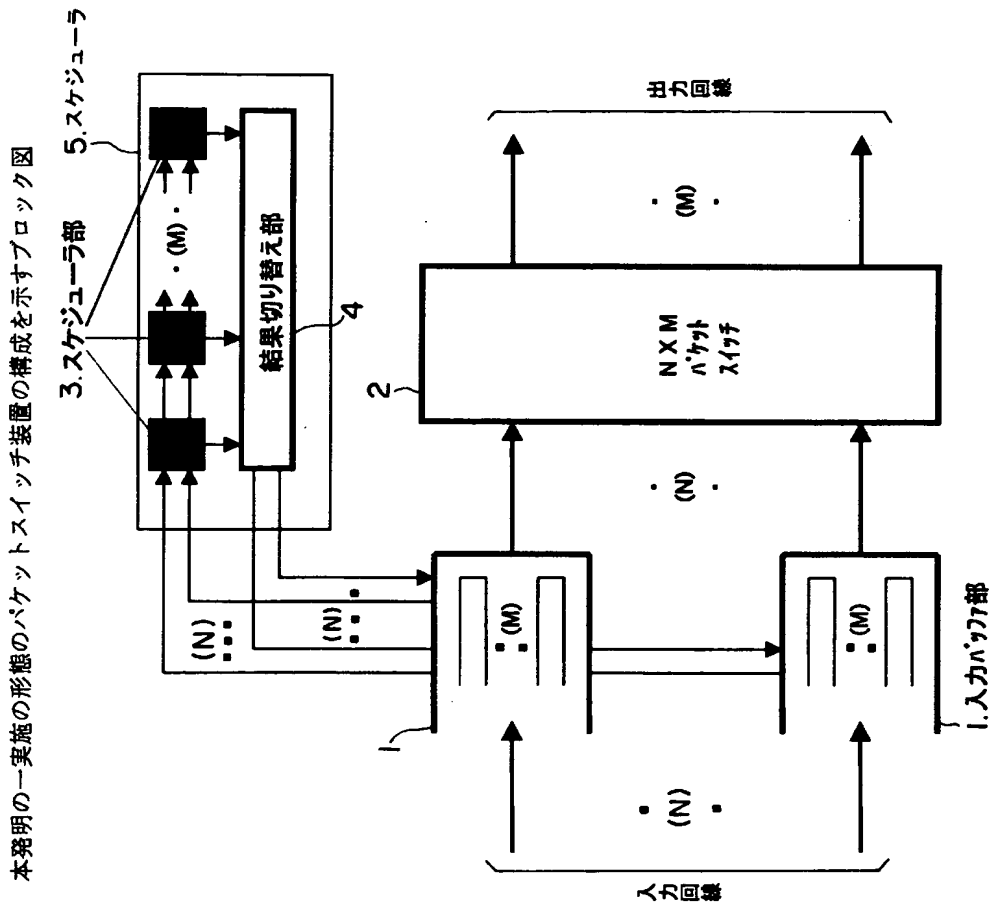
【図 1】

従来のパケットスイッチ装置の構成を示すブロック図

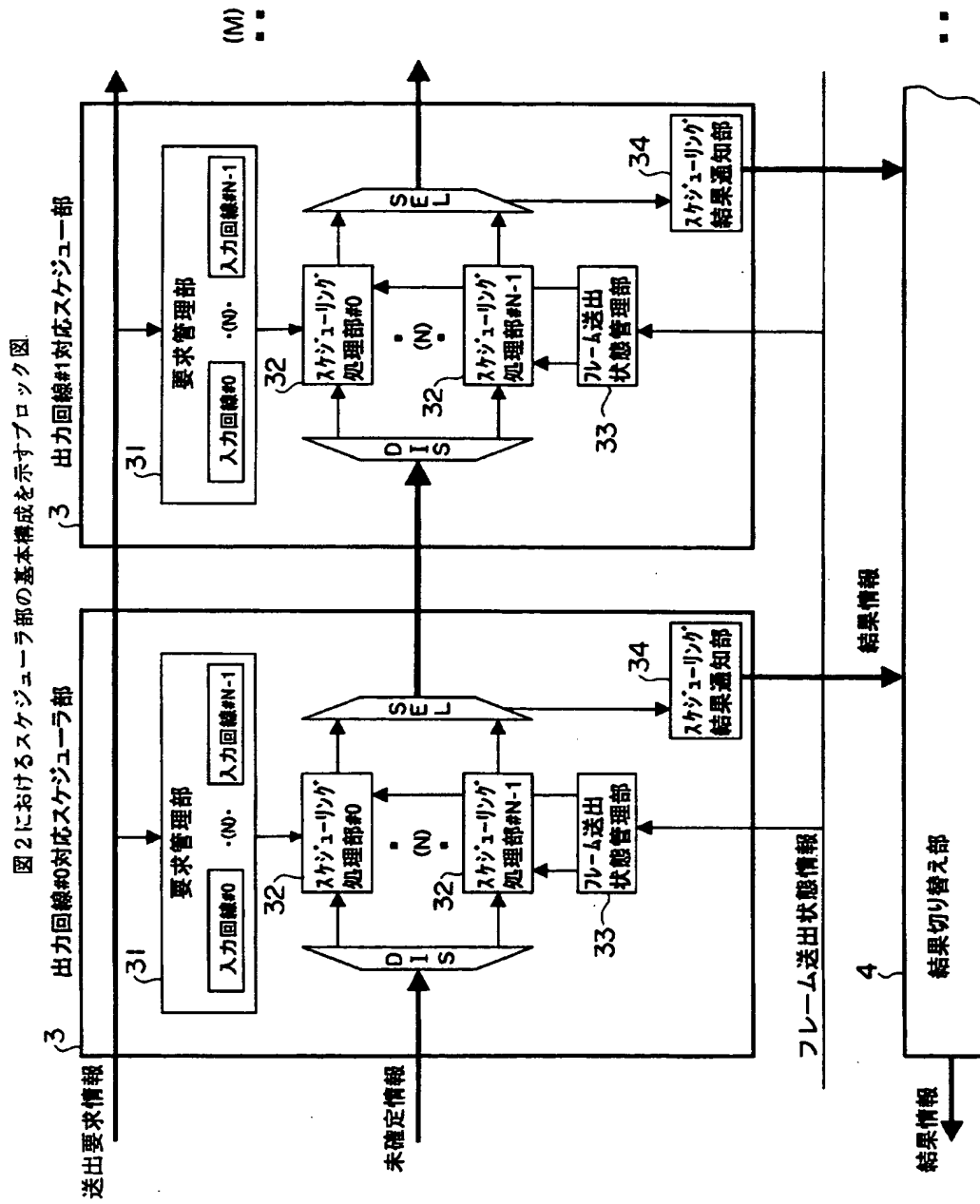




【図 2】

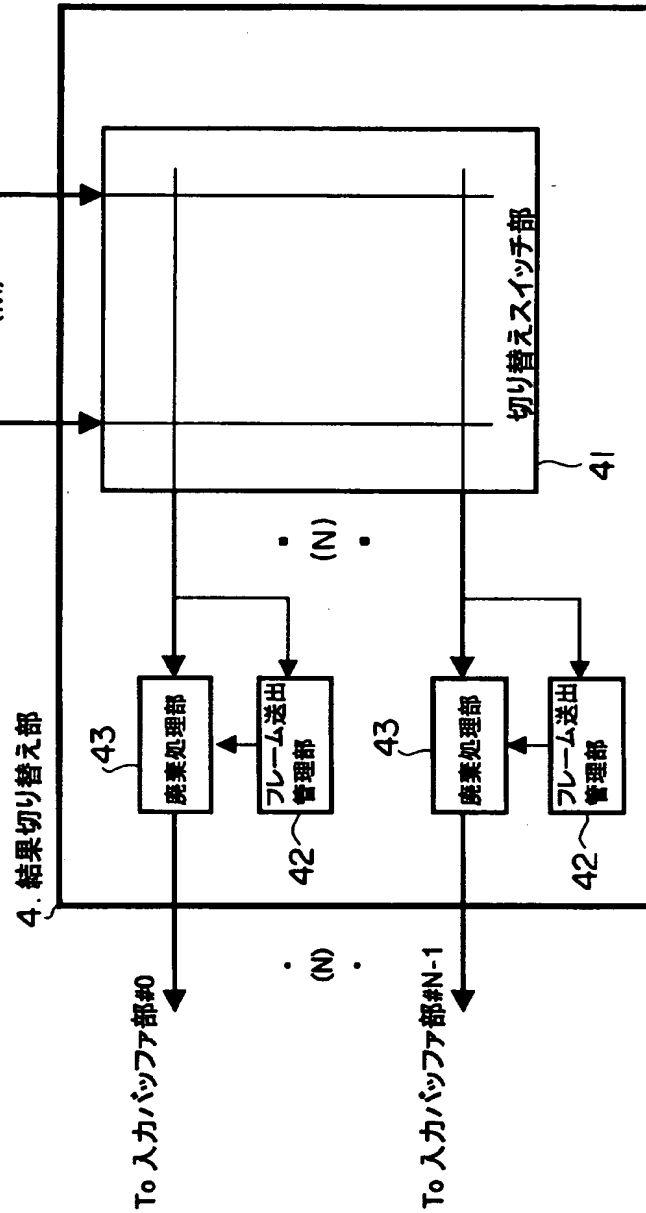


【図 3】



【図 4】

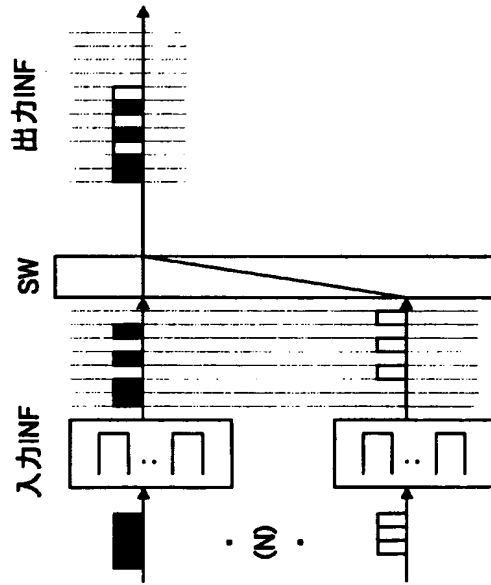
図 2 における結果切り替え部の基本構成を示すブロック図



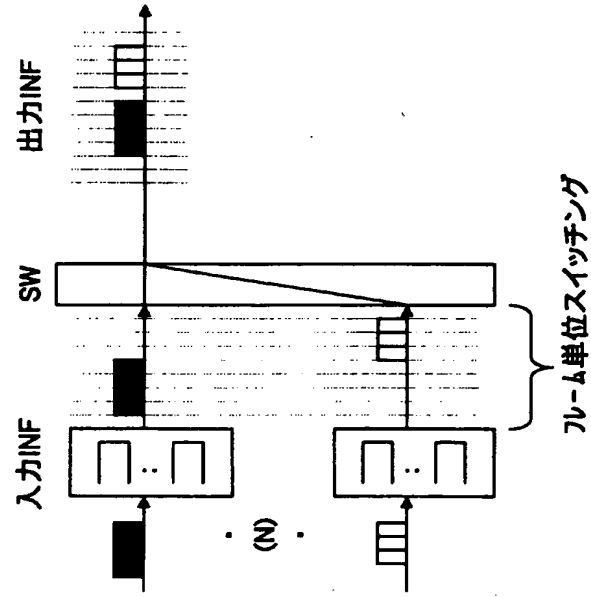
【図 5】

フレーム単位のスケジューリングを説明するための図

(A) フレームを意識しない場合

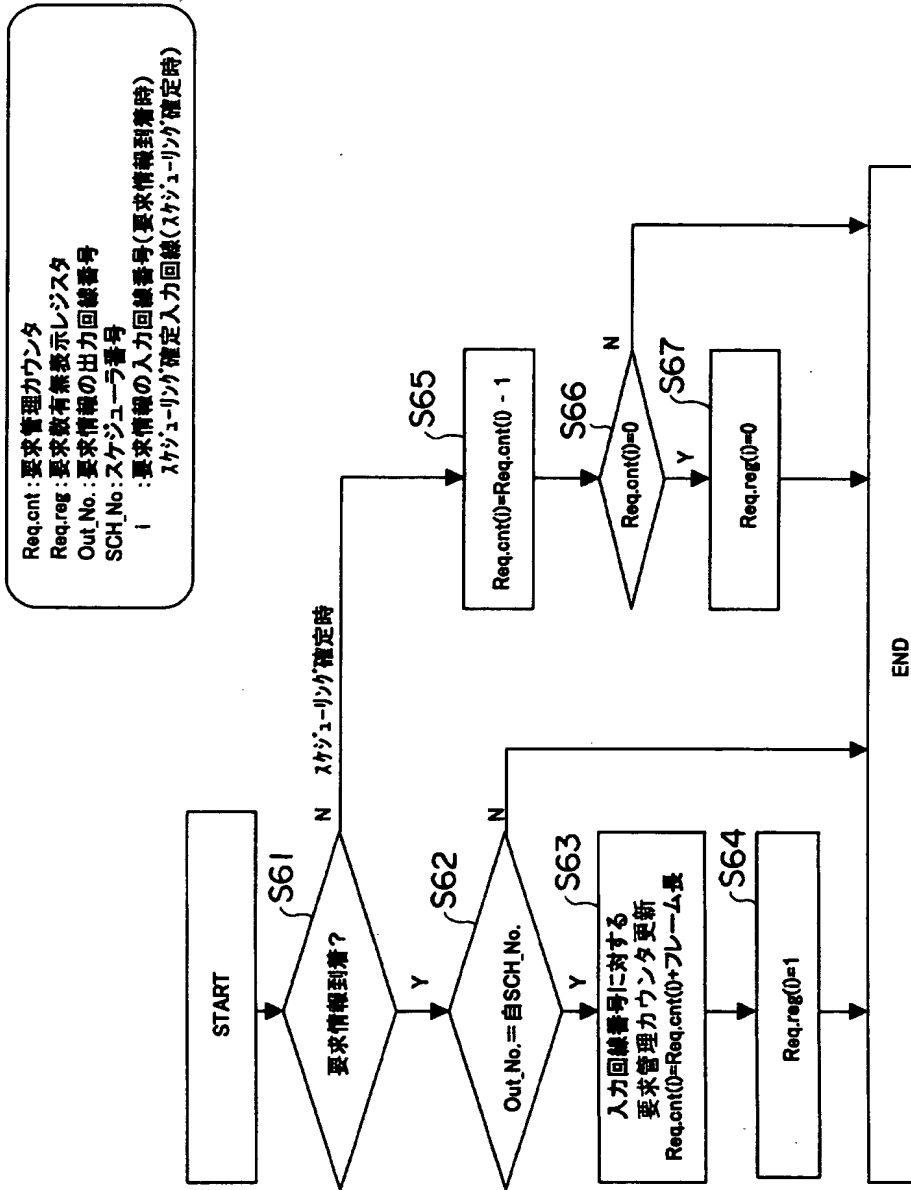


(B) フレームを考慮した場合

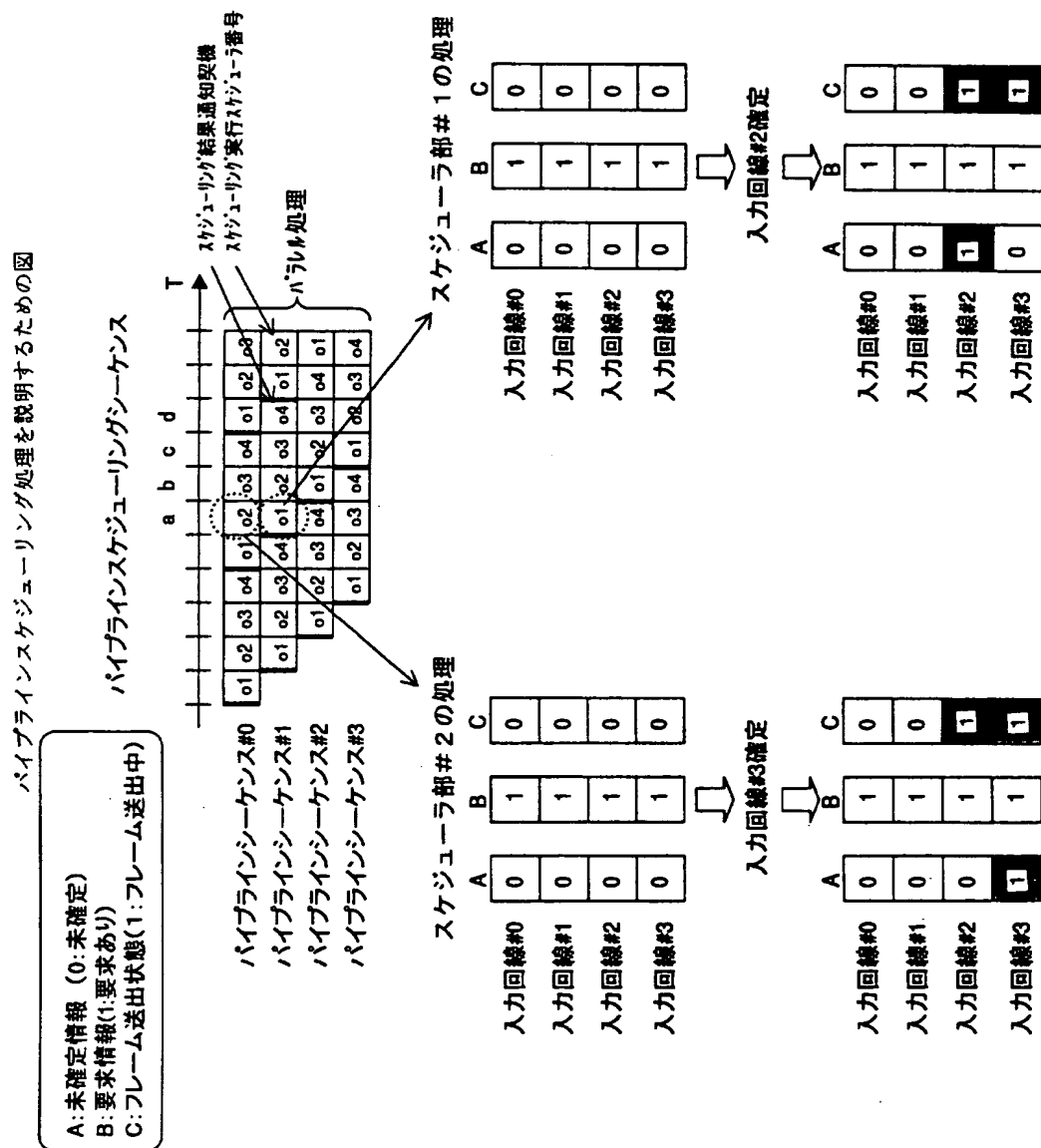


【図 6】

要求管理部の処理手順を説明するためのフローチャート

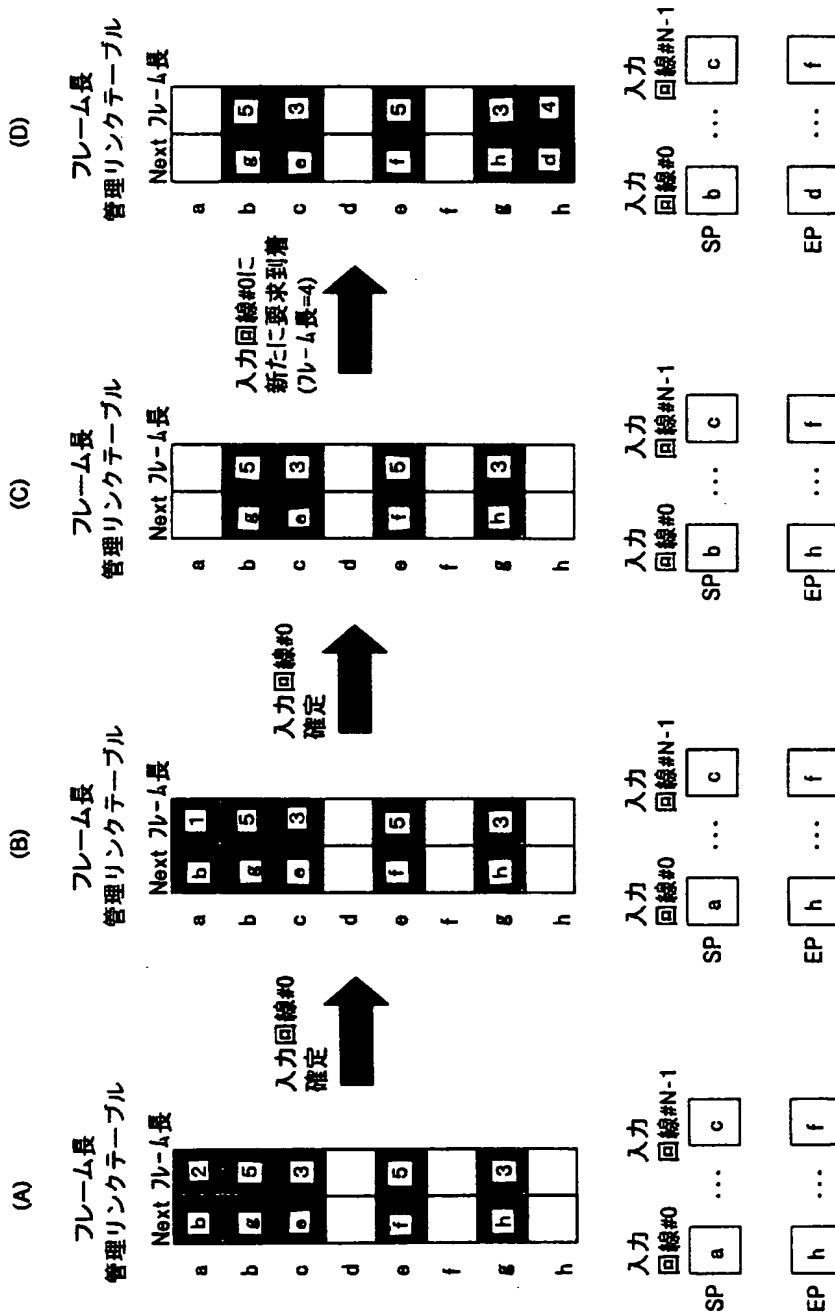


【図 7】



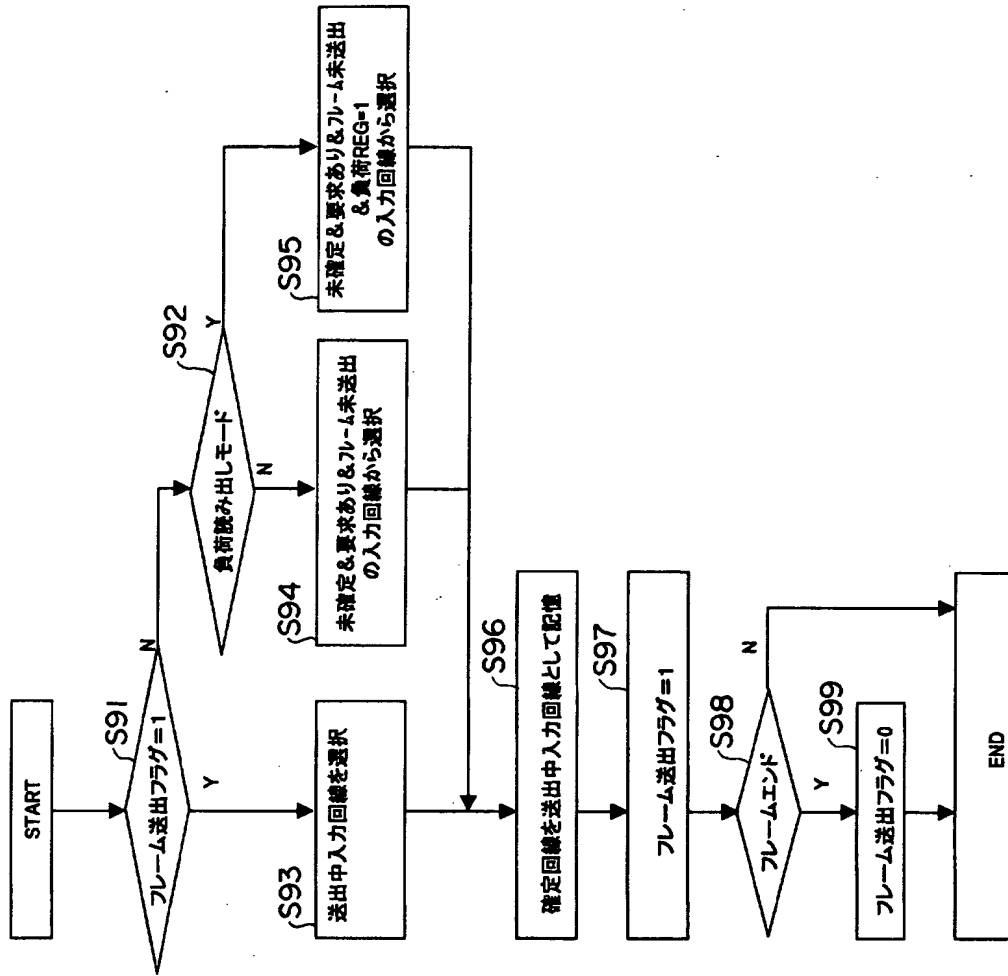
【図 8】

スケジューラ部におけるフレーム長の管理手法を説明するための図



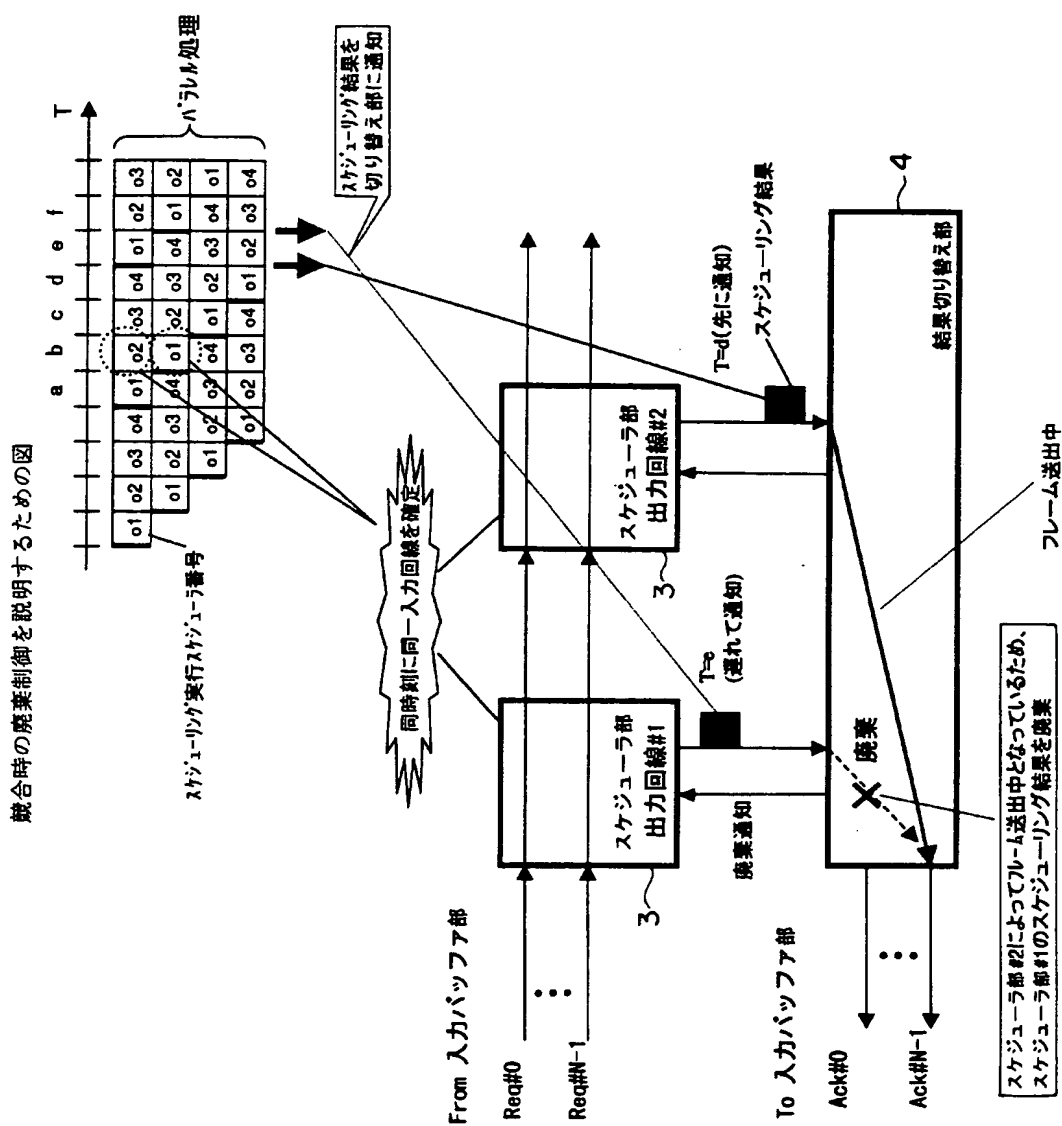
【図 9】

スケジューリング処理手順を説明するためのフローチャート



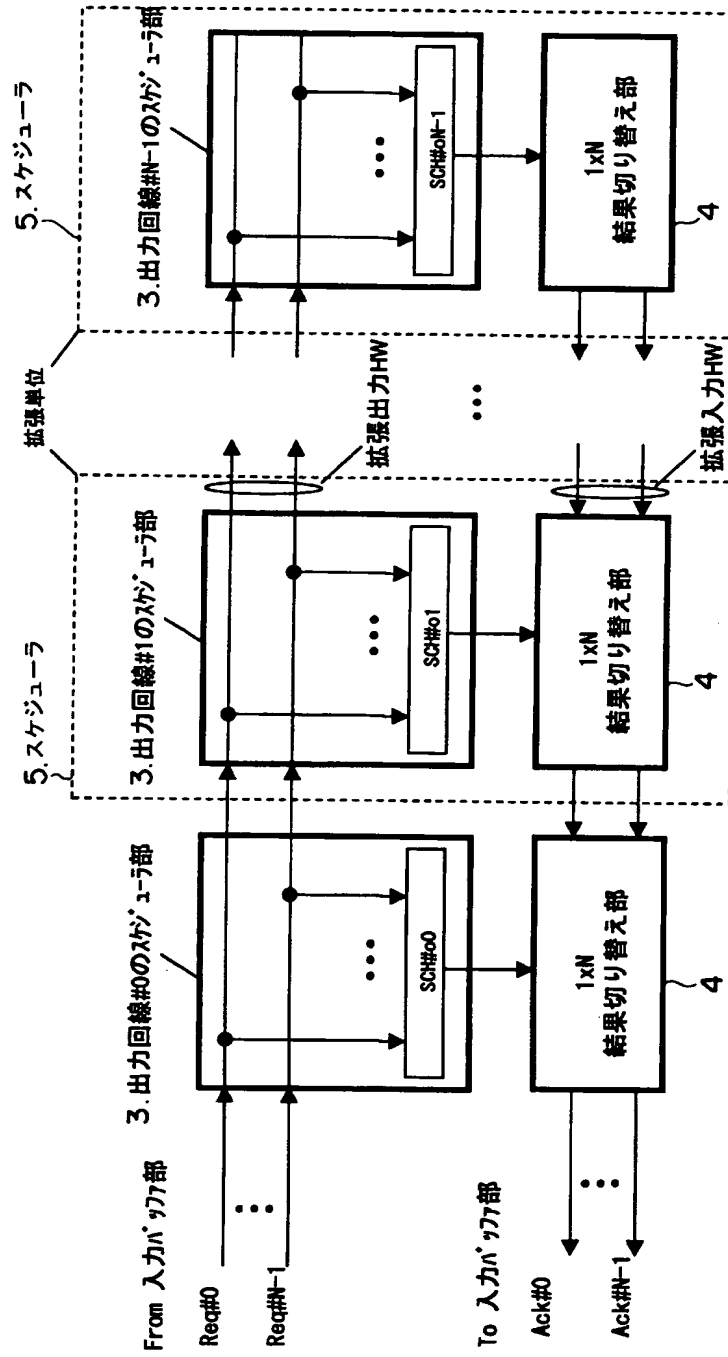


【図 10】



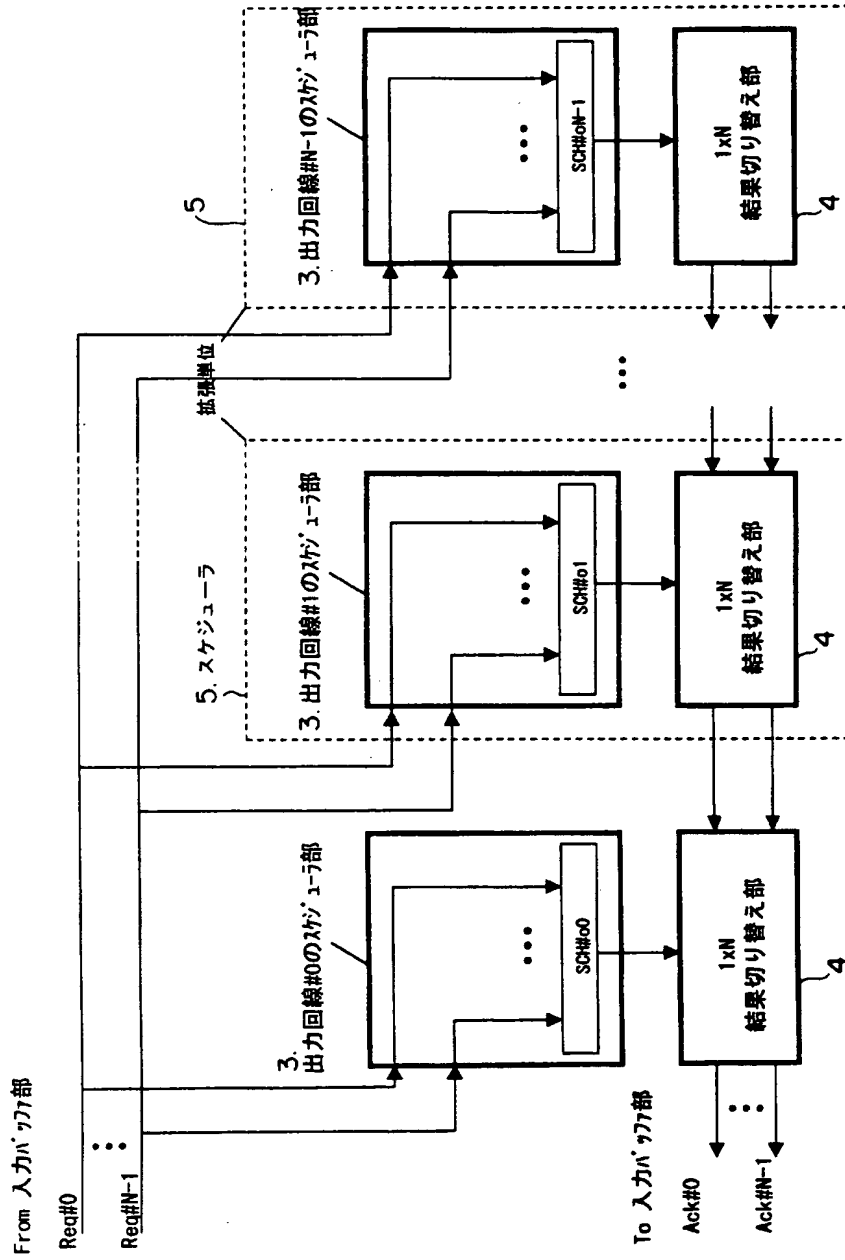
【図 11】

スケジューラの拡張構成の一例を示すブロック図

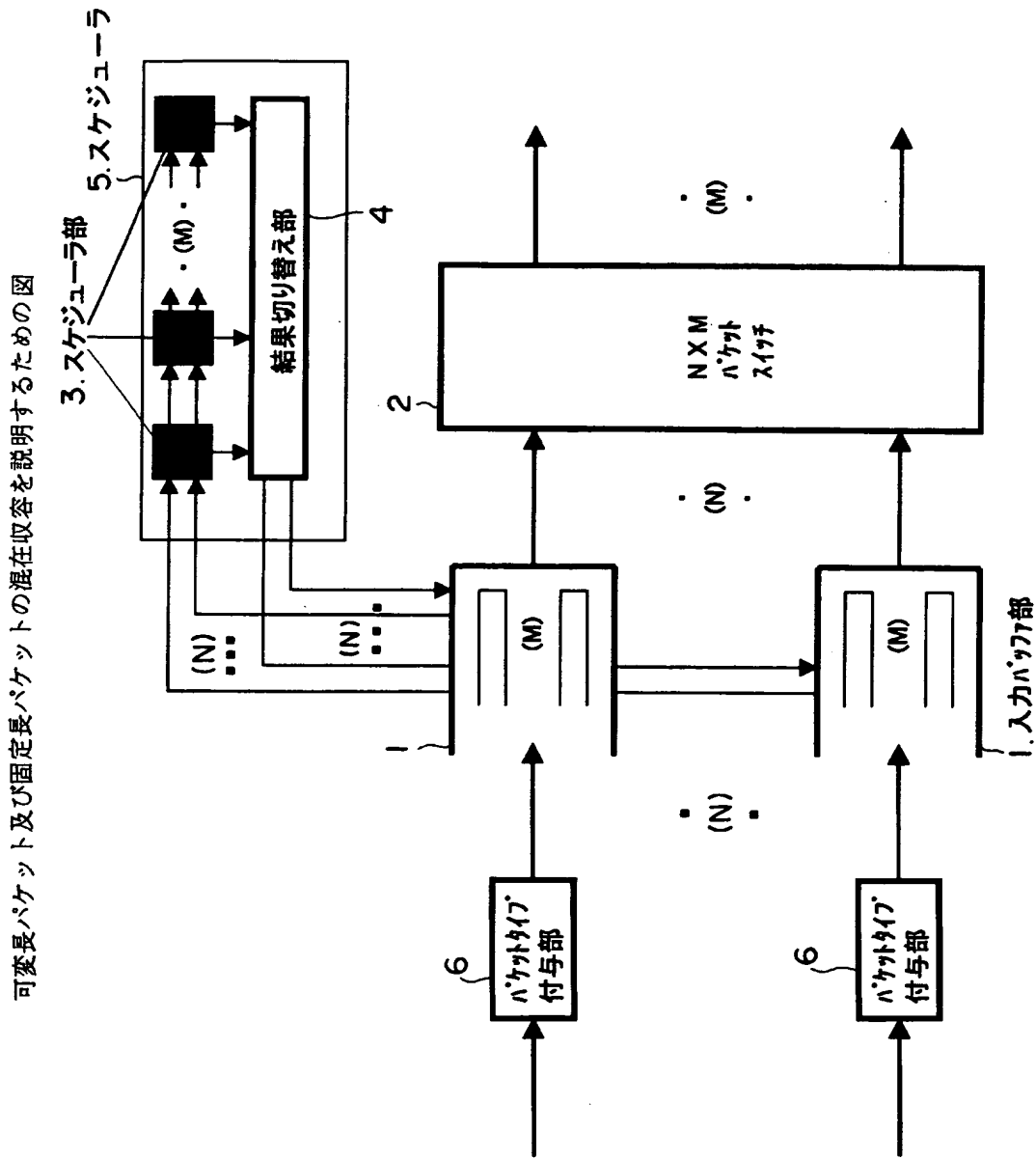


【図 12】

スケジューラの拡張構成の他の例を示すブロック図



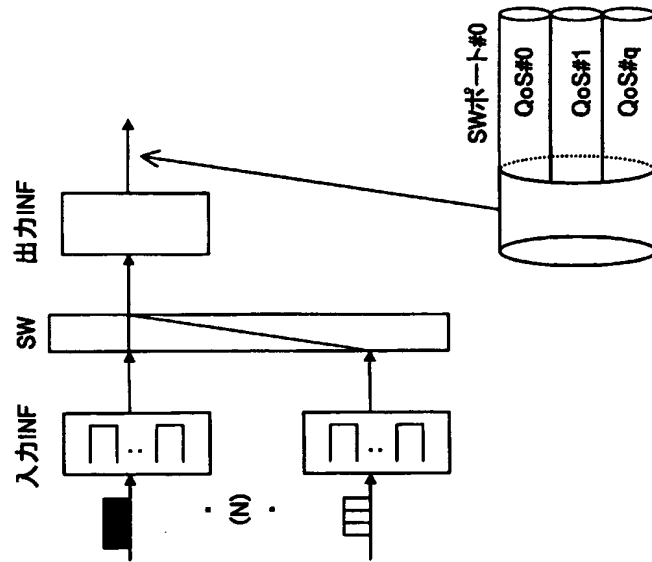
【図 13】



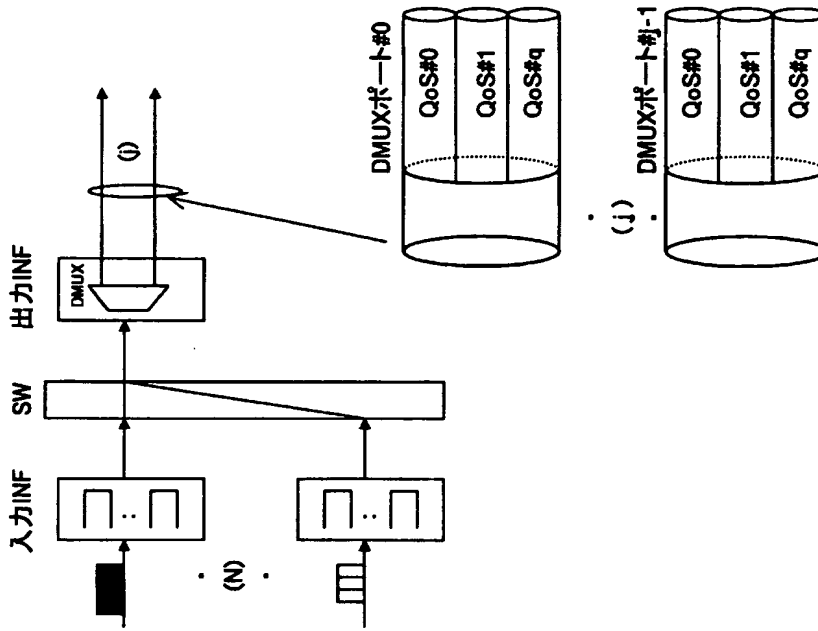
【図 1 4】

QoS制御を説明するための図

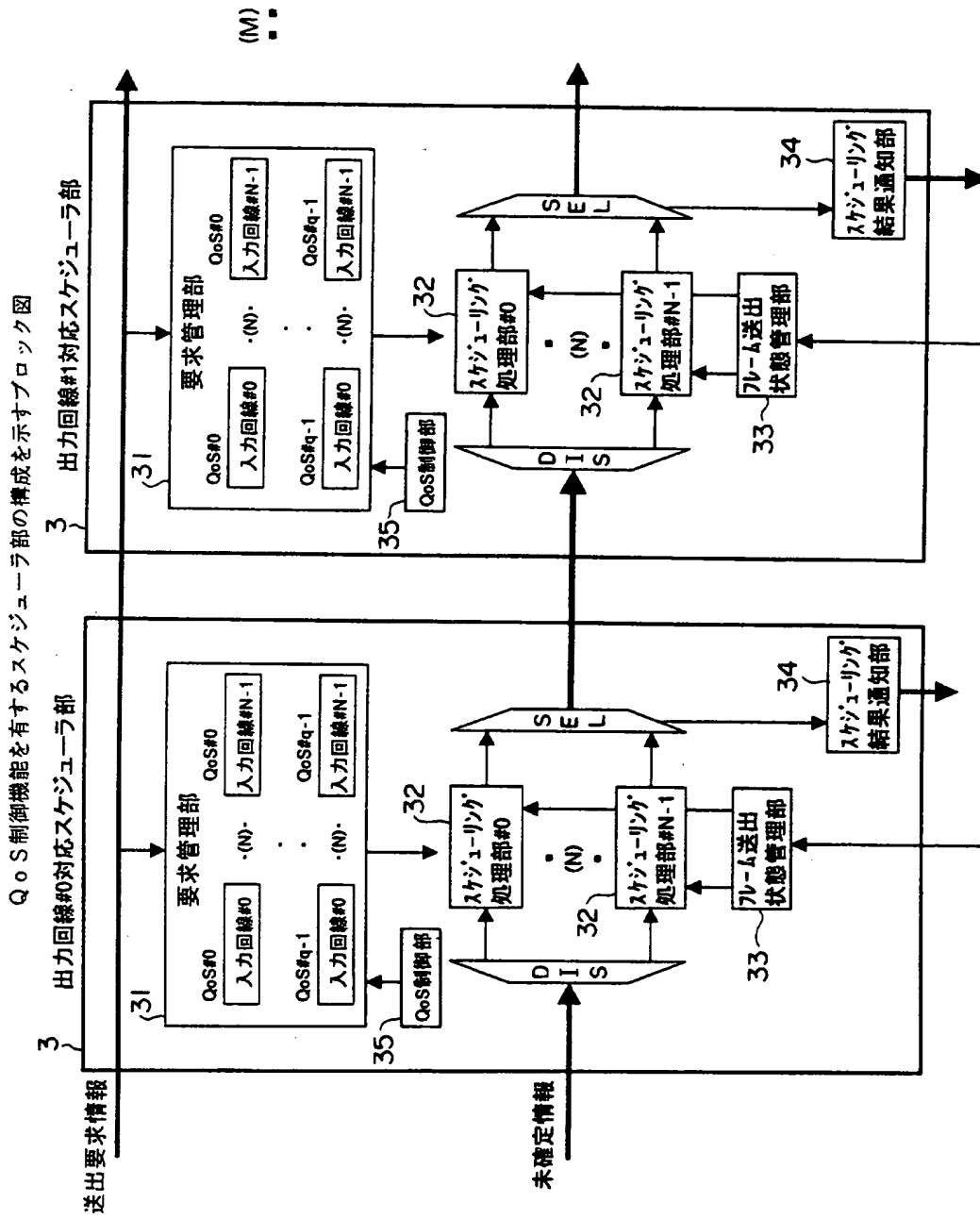
(A) QoS制御構成1



(B) QoS制御構成2(低速回線収容時)

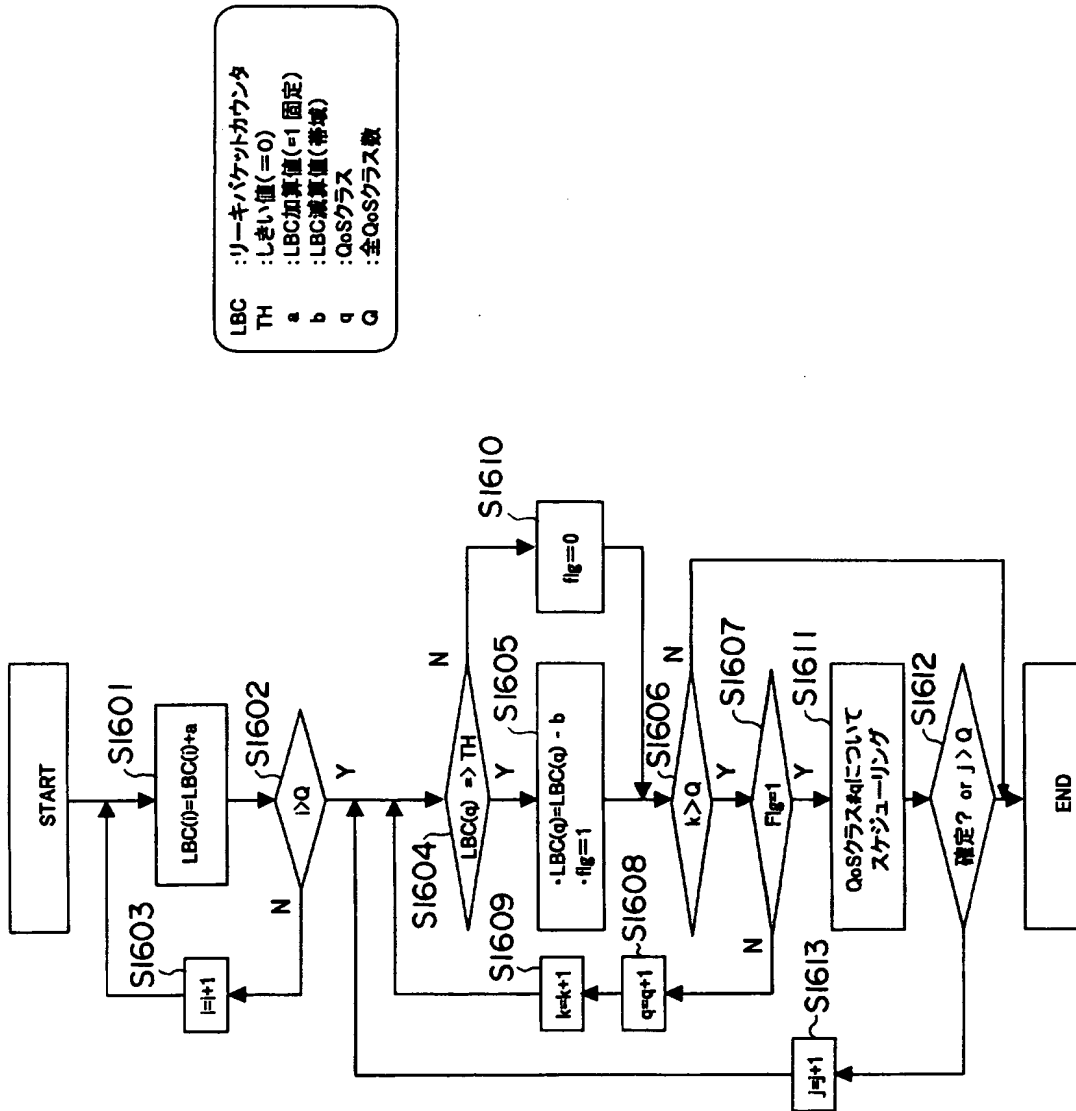


【図 15】

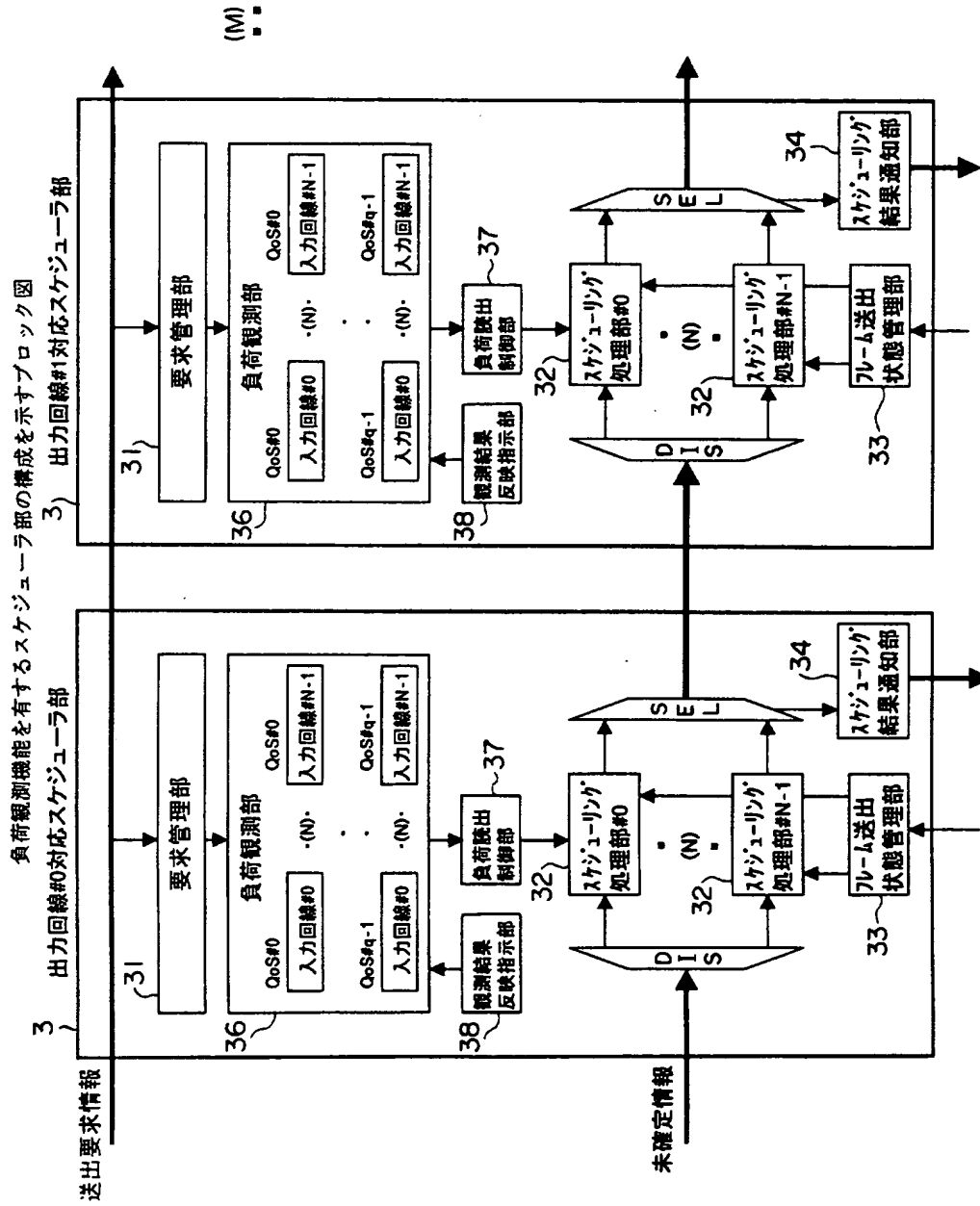


【図 16】

QoS制御の処理手順を説明するためのフローチャート

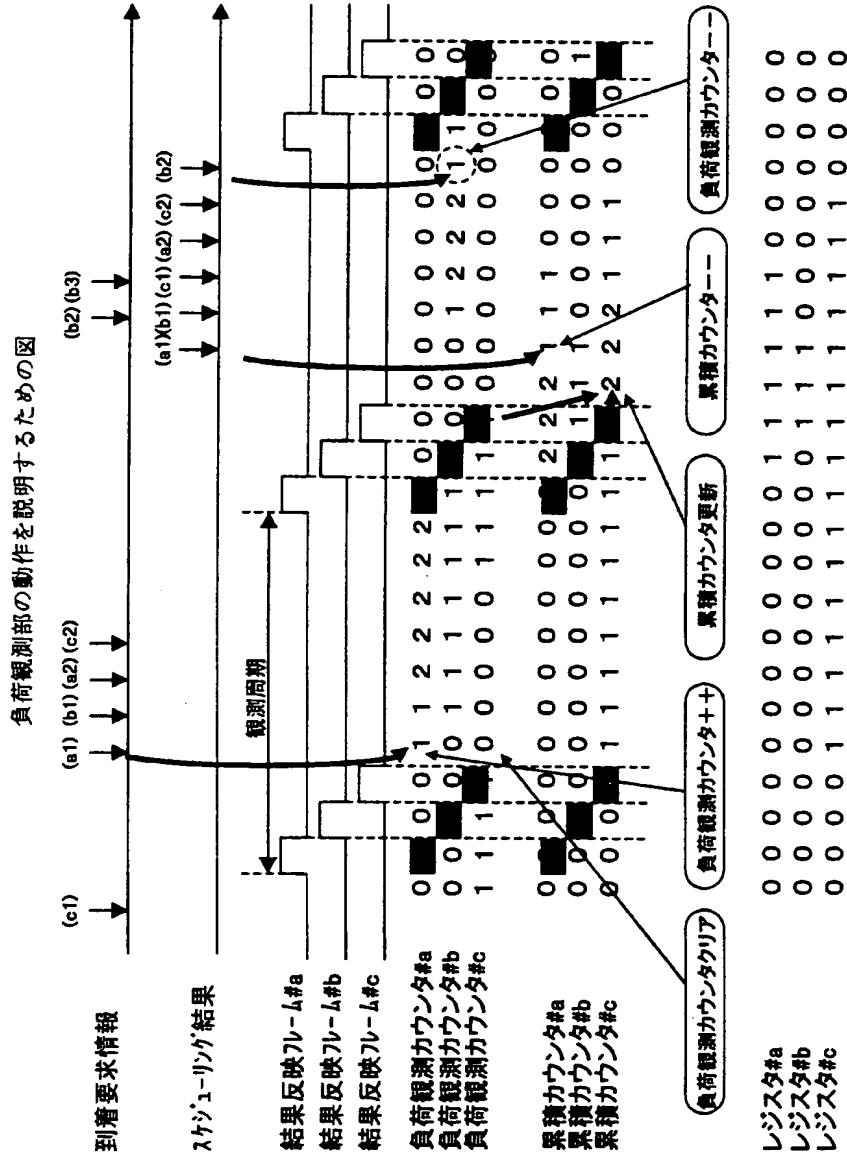


【図 17】

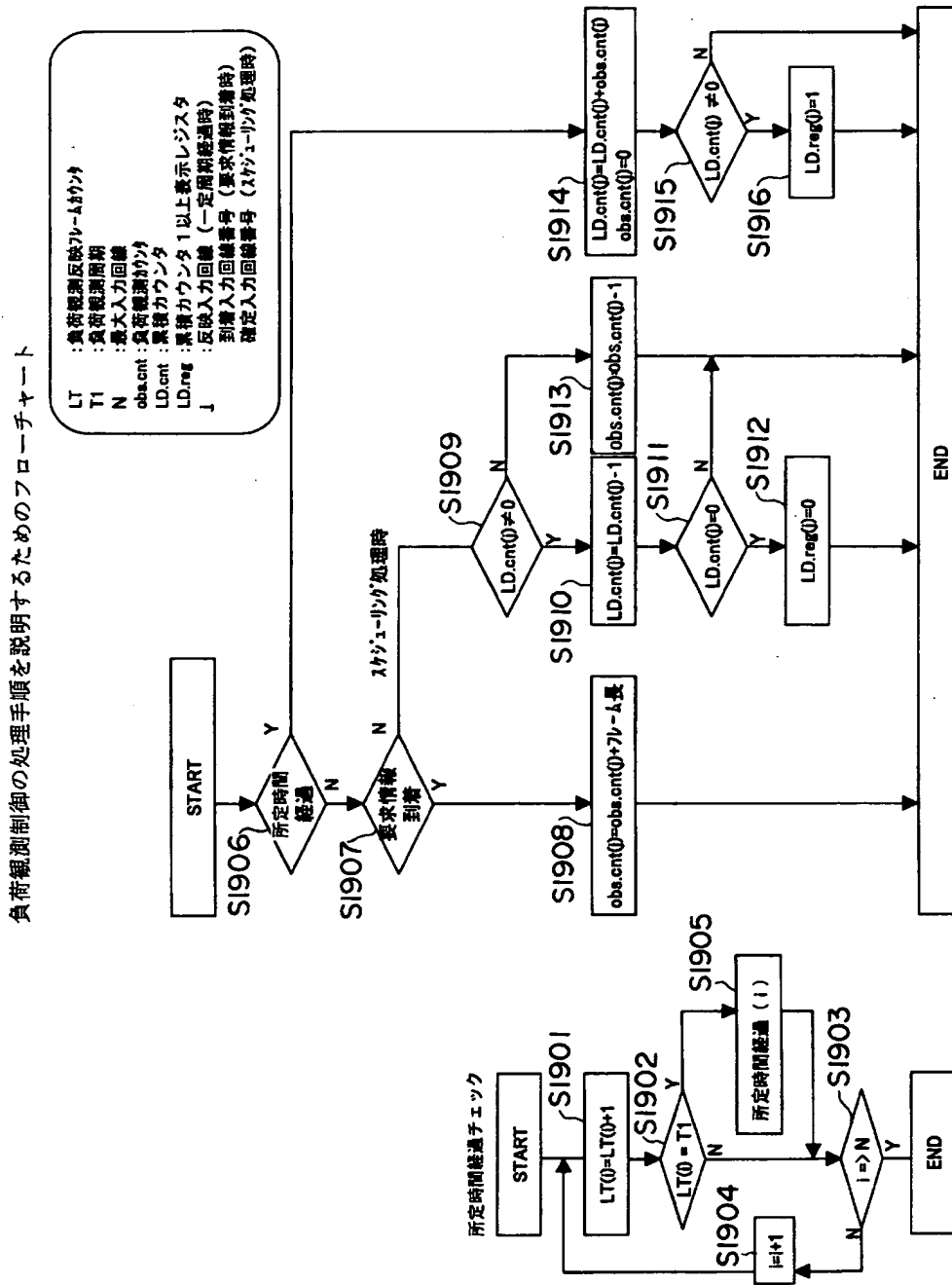




【図 18】



【図 19】



【書類名】 要約書

【要約】

【課題】 固定長パケットのスイッチングを行うパケットスイッチ装置において、IPパケットのような可変長パケットが入力されたときでも効率的に伝送することを可能にする。

【解決手段】 パケットスイッチ装置は、複数の入力バッファ部と、パケットスイッチ部と、前記入力バッファ部からの固定長パケットの送出スケジューリング処理を出力回線の数に対応する複数の単位時間をかけ、かつ前記送出スケジューリング処理を前記入力回線の数に対応する複数並列で実施するパイプライン・スケジューリング処理手段 3 2 と、一つのフレームを構成する前記固定長パケットの送出状態を前記入力回線毎に管理する送出状態管理手段 3 3 とを有し、前記出力回線のいずれかに対応して設けられる複数のスケジューラ部 3 と、前記複数のスケジューラ部のそれぞれにおける前記送出スケジューリング処理の結果情報に対応する前記入力バッファ部に通知する少なくとも一つの結果通知部 4 とを備える。そして、前記複数並列の送出スケジューリング処理において同一フレーム対応の前記固定長パケットを送出中の前記入力回線を選択せず、選択確定後は同一フレーム対応の前記固定長パケットの送出完了まで同一入力回線を選択を維持する。

【選択図】 図 3

出 願 人 履 歴 情 報

識別番号 [000005223]

1. 変更年月日	1996年 3月26日
[変更理由]	住所変更
住 所	神奈川県川崎市中原区上小田中4丁目1番1号
氏 名	富士通株式会社